SEARCHING FOR EXOTIC HIGGS BOSON DECAYS USING A LOW MASS DOUBLE B-TAGGER

BY

MAZIN WOODROW KHADER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Kevin T. Pitts, Chair
Assistant Professor Verena Martinez Outschoorn, Director of Research
Assistant Professor Jessie Shelton
Professor S. Lance Cooper

Abstract

This dissertation presents the development, calibration, and application of a newly developed double b-tagger for use in the context of exotic Higgs decays. After the discovery of the Higgs boson at the Large Hadron Collider, an effort to measure its properties in detail has ensued. Current measurements only constrain non-Standard Model or "exotic" Higgs decays to about $\lesssim 26\%$. Allowing the Higgs to decay into new particles could remedy some outstanding problems in the Standard Model like dark matter or the matter/anti-matter asymmetry in the universe. Thus, an exciting window of opportunity exists for new physics.

The exotic Higgs decays investigated in this work have the Higgs decay into a new spin-zero particle, a. The a-boson then decays into a pair of b-quarks. For low a masses, the b-quarks from the same parent a tend to be collimated. This results in signatures that previous ATLAS reconstruction techniques could not efficiently find. With a newly developed tagger, the $H \to 2a \to 4b$ search in the low mass regime was able to be performed.

To my friends and family.

To both my grandpas.

Rest in peace, Vee-core.

Rest in peace, Sido.

Acknowledgments

The work that is presented in this thesis is the product of collaboration from many wonderful and brilliant people who all deserve more thanks and 1 than what is written here. I first want to thank my advisor Verena Martinez Outschoorn for all the countless hours of help, support, and guidance. She pushed me outside of my comfort zone many times, and it always led me to grow through opportunities I could have never imagined. We took on an ambitious project, but it led to an incredibly rewarding thesis that I can be proud of.

Throughout this endeavor, I have met many mentors and colleagues who have helped me grow as a researcher. I want to thank the post-doctorate Roger Caminal Armadans, who started in the same group around the same time I did. Without his patience and constant source of knowledge helping me ramp up, this thesis would have taken much longer. I also want to thank Ariel Schwartzman and Francesco Rubbo for dedicating so much time coming up with ideas to move this work forward with their quick, smart problem solving mentality. Our major collaborators for the analysis, Aurelio Juste and Javier Montejo Berlingen, deserve a big thanks for providing invaluable feedback and ideas. I also want to thank them for being awesome collaborators that genuinely seek to help and do proper physics, all while having fun. I want to thank Rafael Lopes de Sá for putting in all the countless hours helping on nearly every front of the work described in this thesis. Every question I had would be thoroughly answered and give me an invaluable understanding. Thank you for all the brilliant ideas and explanations. A big thanks to my technical advisor during my qualification task, Andrea Dell'Acqua, who helped me build confidence in myself that has lasted to the end of the thesis. I also want to thank Peter Tornambe for putting in many crunch hours at the end of my thesis to make this work possible. And of course, I also want to thank my thesis committee for providing support, understanding, and time to read this work and allow me to get to this point.

My family was indispensable as a support system. Specifically, my mom, my dad, and my three sisters. I love you all very much and could not get here without you,especially my mom with her constant support.

And last but not least, the wonderful friends I have made in Illinois. There are many times where this work would have been too overwhelming if not for the emotional support and strong bonds you all

have provided. Wooyoung Moon, Harry Mickalide, Chris Zeitler, Michelle Victora, Cristina Schlesier, John Yoritomo, BuseMao, Matt Zhang, Allycia Gariepy, JC Zeng -you have all had a profound impact on my growth in so many ways and have also allowed this work to be possible, thank you.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The Higgs discovery in 2012 by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) was considered a phenomenal testament to the already incredibly successful Standard Model of Particle Physics. However, despite describing most of the fundamental interactions to amazing precision, the Standard Model is known to be incomplete. Gravity, dark matter, and the matter/anti-matter aymmetry in the universe are just some of the known phenomena that the Standard Model does not account for.

A simple and obvious solution is to extend the Standard Model in some way. Since the Higgs discovery is relatively recent, many of its properties have not been measured to high precision. As a result, the Higgs provides an opportunity to look for new physics that could potentially describe some of the deficits the Standard Model currently has. By introducing "exotic" Higgs decays, a new portal linking the Standard Model to new physics can be made.

A simple extension of the Standard Model that can address many of its current outstanding issues has the Higgs decaying into a new spin-zero particle, a. In many of these models, the a particle most likely decays into a pair of b-quarks.

This thesis describes a search for $H \to 2a \to 4b$ using the pp data set at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector during 2015 and 2016. The search focuses on associated production of the Higgs boson with a Z boson, $pp \to ZH$, and uses the leptons from the leptonic decays of the Z boson to trigger on the events. The signature of interest is two leptons from the Z boson decay and multiple jets (sprays of hadrons resulting from partons) which can be identified as originating from b-quarks. In particular, the b-quarks tend to become collimated as the mass of the a gets smaller. Thus, the two pairs of expected b-quarks from the same parent tend to merge and new reconstruction and tagging techniques are required to improve sensitivity in this low mass regime. The work done in this thesis is the full development and application of a new technique, low mass double b-tagging, making it possible to search for this unique exotic Higgs decay signature.

The thesis is organized as follows. Chapter 2 introduces the Standard Model, details about the Higgs, and exotic Higgs decays. The detector is described in Chapter 3. Chapter 4 then goes over the Monte

Carlo simulations used for physics analyses. An overview of the statistical treatment for the work described in the thesis is summarized in chapter 5. Chapter 6 motivates the need for a low mass double b-tagger. Common reconstruction techniques to understand the detector data are covered in chapter 7. The work done on developing, optimizing, and calibrating the double b-tagger follows in chapter 8. Chapter 9 then describes the application of the tagger to the $H \to 2a \to 4b$ physics search and discusses the results. Finally, chapter 10 summarizes and concludes.

Chapter 2

The Standard Model

The Standard Model is a renormalizable, quantum field theory framework that describes most of the known fundamental particles and their interactions. It is invariant under the symmetry group $U(1)_Y \otimes SU(2)_L \otimes SU(3)_c$, which describe three out of the four fundamental forces currently known: the electromagnetic, weak, and strong force. The fundamental particles are comprised of three *generations*, sets of particles differing in mass, of quarks and leptons (fermions) that interact through the various forces (bosons). These interactions are described through the underlying theories in the SM, referred to as sectors. The SM can be summarized by a Lagrangian defined with fields describing the particles and their interactions. To start, the Lagrangian governing the fermions can be expressed by:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.1}$$

where ψ is the fermion field, γ^μ are Dirac matrices, and m is the mass of the fermion. Varying this Lagrangian yields the famous Dirac equation for a relativistic, free particle.

The fundamental interactions seen in nature can be appropriately expressed as *symmetries*, or local transformations of a symmetry group that leave the Lagrangian invariant. In order to preserve gauge invariance, gauge fields are introduced into the Lagrangian.

There is a corresponding gauge field for each generator of a symmetry group. The group $SU(3)_c$ governs the strong force and has 8 generators corresponding to 8 gluons mediating the force between quarks. For the weak and electromagnetic force, $U(1)_Y \otimes SU(2)_L$ gives 4 total generators corresponding to the photon, $W^\pm$- and Z- bosons. The methodology for deriving the associated weak and electromagnetic bosons requires more discussion and will be described in the following section. A summary of the bosons and their interactions can be found in table 2.1, note that the Higgs is also included and will be discussed in section 2.2.

Mediator	Mass [GeV]	Interaction	Electric charge	Spin
Gluon ($\times 8$) (g)	0	strong	0	1
Photon (γ)	0	electromagnetic	0	1
Z	91.19	weak (neutral)	0	1
$W^{\pm}$	80.395	weak (charged)	± 1	1
Higgs (H)	125.7	-	0	0

Table 2.1: Boson classification in the Standard Model.

2.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) describes the strong interaction using $SU(3)_c$. The associated quantum number is referred to as color and can take one of three values (r,b,g). Enforcing that the Lagrangian is invariant under local transformations of this group is achieved by introducing the covariant derivative:

$$D_\mu \equiv \partial_\mu - i g_s T_a G^a_\mu \tag{2.2}$$

where g_s is the strong coupling (commonly shown as $\alpha_s \equiv g_s^2/4\pi$), T_a are the generators of $SU(3)_c$, and G^a_μ is the gluon gauge field. The field strength tensor $G^a_{\mu\nu}$, related to the gauge fields, is also a useful object for constructing gauge invariant quantities that are allowed in the Lagrangian:

$$G^a_{\mu\nu} \equiv \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu \tag{2.3}$$

where f_{abc} are the structure constants of $SU(3)_c$. Thus, the full Lagrangian for QCD and be expressed as:

$$\mathcal{L}_{QCD} = \bar{q}(i\gamma^\mu D_\mu)q - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} \tag{2.4}$$

with q being a 3-component vector for each color for a given quark.

One important feature of the QCD coupling is the particular dependence on the scale of the interaction when first-order quantum corrections are taken into account:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\log(\frac{Q^2}{\lambda^2_{QCD}})} \tag{2.5}$$

where Q is the scale of the interaction (typically characterized by the momentum transfer of the interaction), n_f is the number of relevant quarks ($m_q < Q$), and λ_{QCD} is the QCD scale, an experimentally determined value that gives an approximate divide for when perturbation theory is applicable. The consequence of

equation 2.5 is that the coupling decreases with higher scales and vice versa. This leads to a few important properties of QCD. The first of note is that at higher energies, the coupling becomes smaller and perturbation theory can be applied to computing observables. At infinitely high energies, the coupling goes to zero and quarks will behave like free particles (referred to as asymptotic freedom). If the energy is low, the coupling becomes much stronger and leads to confinement: the restriction that quarks and gluons cannot be free.

The consequence of these properties is that as quarks or gluons separate from each other, it eventually becomes more energetically favorable to produce a quark-antiquark pair from the vacuum. As quarks and gluons are produced from the pp collisions, this process will happen as they traverse from the interaction point. This leads to a spray of QCD bound states (hadrons). This spray of hadrons is known as a jet and is the signature from a produced quark or gluon in the collider.

2.2 Electroweak Sector

The next area interest in the SM is the electroweak sector, which is the combination of the weak and electromagnetic force. Historically, these forces were considered completely independent with quantum electrodynamics (QED) being developed separately from the weak force (Fermi's theory of beta decay). However, due to Fermi's interaction not being renormalizable and the discovery of massive bosons ($W^{\pm}$, Z) mediating the weak force [1], a new theory was developed that unified the two forces under the group $SU(2)_L \otimes U(1)_Y$. To have a Lagrangian that is invariant under local transformations of this group, another covariant derivative can be introduced:

$$D_\mu \equiv \partial_\mu - igT_iW_\mu^i - ig'\frac{Y}{2}B_\mu \tag{2.6}$$

where g and g' are the couplings of $SU(2)_L$ and $U(1)_Y$ respectively, T_i are the generators of $SU(2)_L$, Y is the quantum number of $U(1)_Y$ (commonly referred to as the hypercharge), and W_μ^i/B_μ are the gauge fields of the respective groups. As before with QCD, gauge invariant quantities can be introduced using field strength tensors for each group:

$$W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon_{ijk}W_\mu^j W_\nu^k B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.7}$$

with ϵ_{ijk} being the antisymmetric Lev-Cevita tensor. The electroweak Lagrangian can be expressed as:

$$\mathcal{L}_{EW} = \bar{f}(i\gamma^\mu D_\mu)f - \frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{2.8}$$

with f being the fermions. It should be noted the only *left-handed* fermions interact via the weak force and transform as $SU(2)_L$ doublets, while right-handed fermions transform as singlets (no interaction).

If mass terms are introduced for the gauge fields or fermions, the symmetry is broken and the Lagrangian is no longer invariant. In order to keep the SM gauge invariant, renormalizable, and remain compatible with experimentally verified massive vector bosons, a mechanism was proposed to modify the EW theory.

2.3 The Higgs Mechanism

The Higgs mechanism is such a process that allows all the aforementioned properties to coexist. It starts with introducing a new weak isospin doublet of complex scalar fields, called the Higgs field:

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.9}$$

where + or 0 refer to the electric charge of the field given by a combination of $SU(2)_L$ and $U(1)_Y$ quantum numbers: $\hat{Q} = \hat{T_3} + \frac{\hat{Y}}{2}$.

The Lagrangian for the Higgs field can be comprised of gauge invariant terms as before:

$$\mathcal{L}_\Phi = (D_\mu \Phi)^\dagger (D^\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \tag{2.10}$$

where D_μ is the same covariant derivative as equation 2.6. If the potential is minimized, this leads to a set of solutions that depends on the signs of the parameters μ^2 and λ. When λ is negative, this leads to unphysical and unstable equilibria. When λ is positive, the solutions are either 0 for positive μ^2 or

$$\Phi^\dagger \Phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v}{2} \tag{2.11}$$

for negative μ^2, which is non-zero, degenerate solution.

When the system settles into a particular ground state, the symmetry $SU(2)_L \otimes U(1)_Y$ is broken. This is referred to as spontaneous symmetry breaking (SSB). Certain systems exhibit a process where a system settles into a ground state from a degenerate set. The symmetry is not broken in the Lagrangian, but the symmetry is not respected when the system chooses a given ground state.

According to Goldstone's theorem, massless scalar bosons show up if a continuous symmetry is spontaneously broken. If the system and Higgs field are analyzed as a perturbation around the ground state, Φ

can be expressed as:

$$\Phi(x,t) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ (v + h(x,t)) \end{pmatrix} e^{i\vec{\sigma}\cdot\vec{\xi}(x,t)} \tag{2.12}$$

where h and $\vec{\xi}$ are real scalar fields.

It turns out that the ξ fields can be set to zero due to gauge invariance (the particular gauge choice is referred to as the unitary gauge). Using 2.12 in 2.6 after applying the gauge and taking the relevant result yields:

$$\begin{aligned} &\left| \left(-ig\frac{\sigma}{2}\vec{W}_\mu - i\frac{g'}{2}B_\mu \right) \Phi \right|^2 \\ &= \frac{1}{8} \left| \begin{pmatrix} gW_\mu^3 + g'B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -gW_\mu^3 + g'B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} \right|^2 \\ &= \frac{1}{8}v^2g^2 \left[(W_\mu^1)^2 + (W_\mu^2)^2 \right] + \frac{1}{8}v^2(g'B_\mu - gW_\mu^3)(g'B^\mu - gW^{3\mu}) \end{aligned} \tag{2.13}$$

With a bit of algebra, the terms can be expressed in terms of mass eigenstates:

$$W_\mu^\pm = (W_\mu^1 \mp iW_\mu^2)/\sqrt{2} \tag{2.14}$$

$$Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{(g^2 + g'^2)}} \tag{2.15}$$

$$A_\mu = \frac{g'W_\mu^3 + gB_\mu}{\sqrt{(g^2 + g'^2)}} \tag{2.16}$$

which are the $W^\pm/Z$ bosons and the photon with masses $m_W = \frac{vg}{2}$, $m_Z = v\frac{\sqrt{g^2+g'^2}}{2}$, and $m_\gamma = 0$, respectively. The degrees of freedom that were in the original Higgs field Φ are now in the massive gauge bosons (1 additional degree of freedom for each previous massless boson that gained a mass term).

There are no restrictions preventing the Higgs field from coupling to the fermions. A gauge invariant term can be introduced to the SM Lagrangian:

$$\mathcal{L}_{YW} = \sum_{f=l,q} \lambda_f \left[\bar{f}_L \Phi f_R + \bar{f}_R \bar{\Phi} f_L \right], \tag{2.17}$$

where the matrices λ_f describe the Yukawa couplings between the single Higgs doublet and the fermions. Using equation 2.12 in 2.17 results in interaction terms between the physical Higgs field and fermions, which allows mass terms for the fermions.

2.4 Properties of the Higgs

Since the work of this thesis focuses on looking for new physics through the Higgs boson, this section will showcase some of the important and relevant properties of the Higgs.

2.4.1 Production Modes from Proton-Proton Collisions

The main production mechanisms of the Higgs from proton-proton collisions are gluon fusion (ggF), weak vector boson fusion (VBF), associated production with a weak vector boson ($ZH/W^{\pm}$), and associated production with top/anti-top quark pairs ($t\bar{t}H$). Feynman diagrams for the processes can be seen in 2.1. The theoretically computed cross-sections for each production mode as a function of the center-of-mass energy $\sqrt{s}$ can be seen in Fig. 2.2, with (N)NLO/NNLL referring to higher order corrections in the computation in the cross-section (described in more detail in the following chapter).

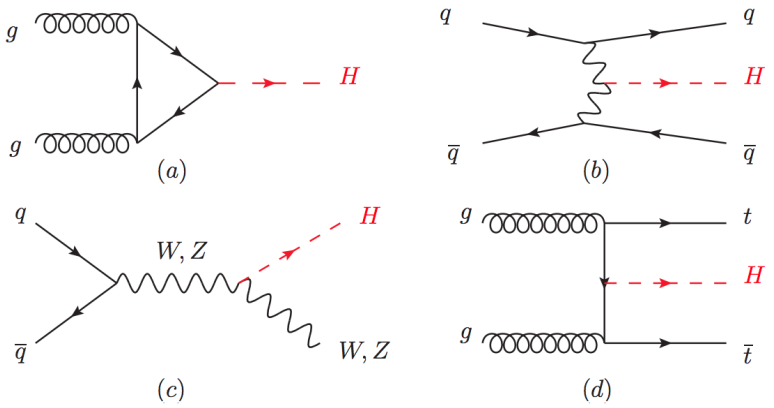

Figure 2.1: Feynman diagrams for the four primary production modes ggF (a), VBF (b), $ZH/W^{\pm}H$ (c), and $t\bar{t}H$ (d).

2.4.1.1 Gluon Fusion

The production mode with the highest cross-section through pp collisions is through gluon-gluon fusion. As shown in Fig. 2.1, the gluons fuse into a Higgs mediated by a quark loop. Since the lowest order computation for the process depends on the mass squared of the quark in the loop, the quark with highest mass dominates the contribution, i.e. the top quark. Although loop contributions tend to be suppressed, hadron colliders have enormous contributions from gluon collisions, resulting in ggF being the dominant production mode. The final signature from ggF is simply the Higgs's decay products, which will be described in 2.4.2 (any

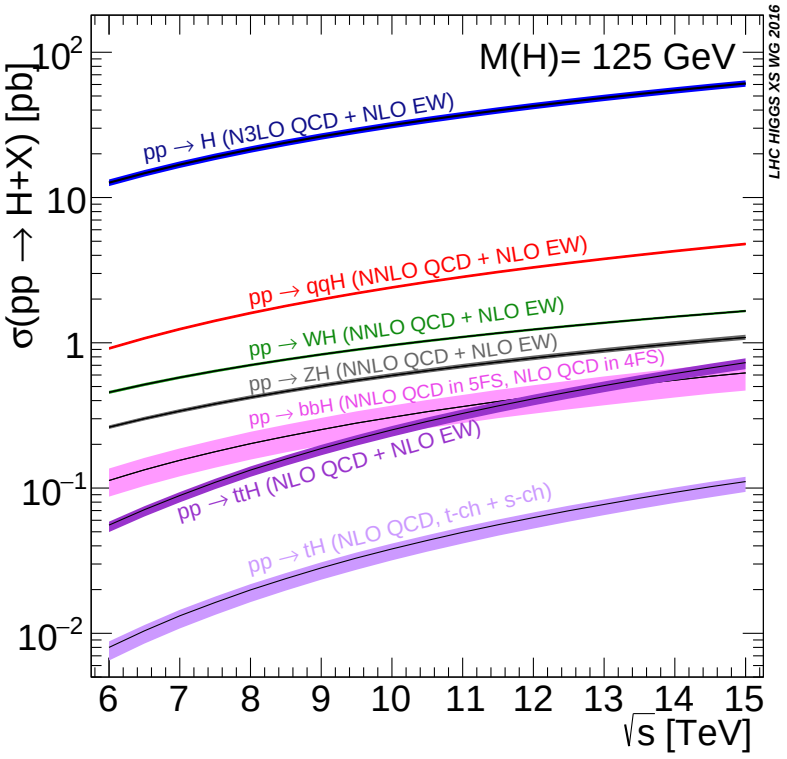

Figure 2.2: Theoretical cross-sections for each production mode as a function of the center-of-mass energy $\sqrt{s}$. (N)NLO/NNLL refer to higher order corrections in the computation in the cross-section.

initial or final state radiation will also contribute to the final signature, but that is true for all production modes).

2.4.1.2 Vector Boson Fusion

The next highest cross-section process is weak vector boson fusion. The Higgs is produced via radiation from a Z or $W^{\pm}$ boson propagator that exchanges between two scattering quarks. These quarks tend to leave signatures in forward regions of detectors, where the angle with the beamline is small. Additionally, the Z and $W^{\pm}$ do not directly interact with gluons and so radiation is suppressed in the central region for detectors (higher angle from the beamline). These two attributes give VBF a unique signature to distinguish against the myriad QCD processes produced in a hadron collider.

2.4.1.3 Associated Production with a Weak Vector Boson

Following VBF by order in cross-section is $Z/W^{\pm}$ associated production. Two quarks can fuse together to create a weak vector boson that radiates a Higgs. Both Z and $W^{\pm}$ bosons can decay into leptons a

reasonable amount of the time (10% and 30% respectively). Leptons (specifically electrons and muons) tend to leave reasonably clean signatures in hadron detectors and thus provide a good handle for identifying potential Higgs events. The leptons also provide good trigger objects and reduce backgrounds. Because of the increased sensitivity, the ZH production mode is the process probed for the work done in this thesis.

2.4.1.4 Associated Production with Top-Quark Pairs

The last major production channel comes from $t\bar{t}H$. The Higgs is produced via radiation through a top quark pair. Each of the top quarks decay into a b-quark and $W^{\pm}$ boson. Leptonic decays of the $W^{\pm}$ can provide clean signatures, along with the b-quarks, at the cost of a low cross-section.

2.4.2 Decay Modes of the Higgs Boson

The dominant modes have the Higgs decaying into a bottom quark pair ($b\bar{b}$) or two W bosons where one is off-shell. The following subdominant decays are into a pair of gluons, τ leptons, charm quarks, Z bosons (one off-shell), photons (γ), or muons. The theoretical computation for the Higgs decay modes in terms of their branching ratio (BR) can be seen in Fig. 2.3. The viable final signatures for detecting Higgs events depends on both the production and decay modes. For example, if one wanted to measure the b-quark pair decay channel, trying to find signatures resulting from ggF production runs into practical issues. The combination may have the highest cross-section and BR, but the final signature results in purely hadronic signatures which make it difficult to distinguish from the large, similar background processes produced at a hadron collider. Often times, a more optimal combination exists. For the example with the Higgs decaying into a hadronic final state, $ZH/W^{\pm}H$ or $t\bar{t}H$ are considered to be more viable due to the additional leptons in the final state.

Aside from measuring prediction yields directly, another important experimental measure for the decay modes is the decay width. The total width is just the simple sum of the widths of the individual decay modes and any constraints on the measured mass width correspond to constraints on the decay modes. The SM Higgs mass width is predicted to be around $\Gamma_h \approx 4.07\ MeV$.

2.4.3 Current Status of Higgs Measurements

The Higgs boson was discovered by observing a small number of decay modes [2, 3]. The decay modes measured in ATLAS, shown in Fig. 2.4, are consistent with the SM predictions, but they are measured with limited precision and constitute only a subset of the Higgs boson decay modes.

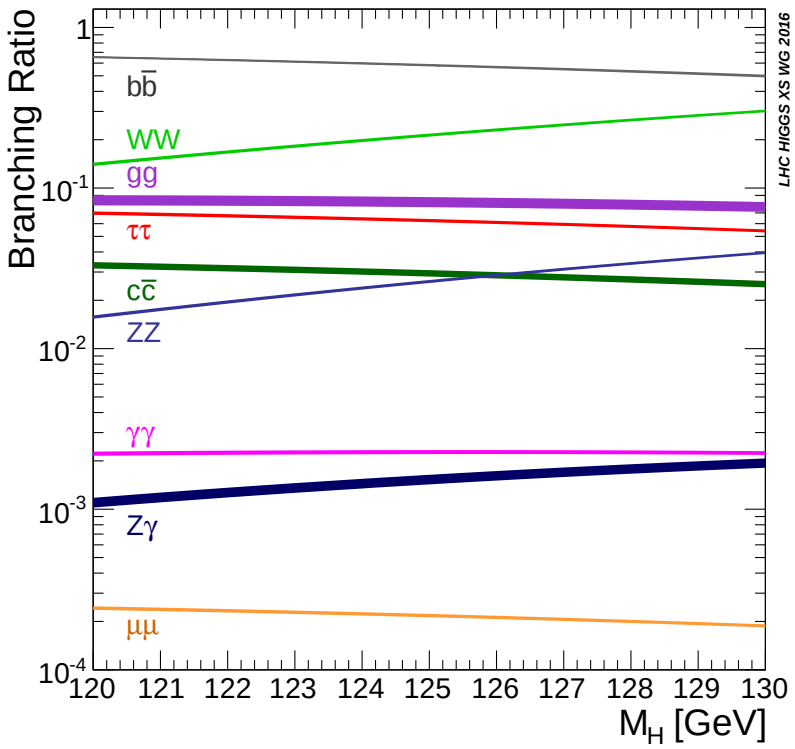

Figure 2.3: Theoretical branching ratios for the Higgs boson as a function of its mass.

As a result, available measurements are only able to constrain non-SM decays to $\lesssim 26\%$ (depending on the model) of all decays at 95% confidence level (C.L.) [4–7]. Even with the full data set (3000 fb^{-1} of data) from planned future program, HL-LHC, the expected precision to measure SM decays is $\sim 10\%$ [8]. Thus, for the lifetime of the LHC, there will still be room for new physics in Higgs boson decays at least at the $\sim 10\%$ level. These decays are interesting because they may be a window to new physics, including scenarios called *Higgs portals*, where the Higgs boson is the leading or only mediator between the SM and new physics sectors. These exotic decays can alleviate the deficits of the SM, which are described in the next section.

2.5 Shortcomings of the Standard Model

Despite the myriad successful predictions from of the SM measured with high precision, there are still major phenomena completely unexplained by the SM. Gravity, one of the four known fundamental forces, is not included in the SM. A discrepancy between the expected and observed gravitational behavior of galaxies based on the observed matter led to the inference of a non-luminous substance, dark matter. The SM has no candidate particles to explain dark matter, which is estimated to comprise around 80% of the matter content

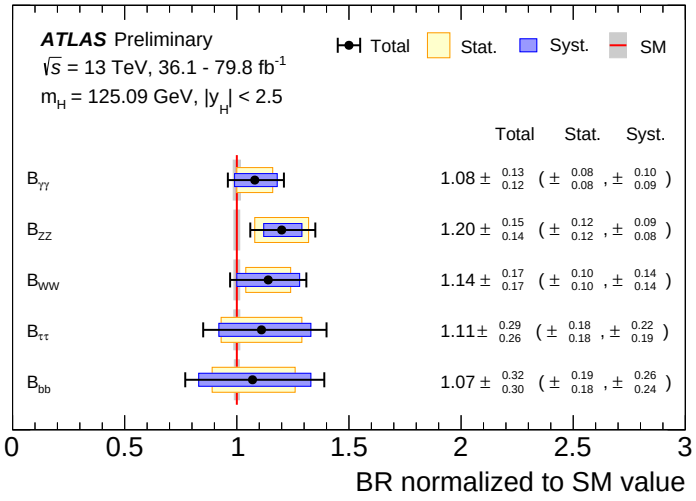

Figure 2.4: Ratio of the production cross section times branching ratio that best fits available data relative to the SM expectation for individual decay modes [7].

in the universe. The apparent matter/anti-matter asymmetry in the universe is also not adequetely explained in the SM. The experimental observations of neutrino oscillations imply that neutrinos have differences in mass, which are not currently supported by the SM.

Another important issue is the hierarchy problem. Quantum corrections to the Higgs mass result in quadratically divergent terms that are sensitive to the scale of the interactions. This especially becomes problematic if the SM is supposed to be valid up until Planck scales, leading to corrections of 30 orders of magnitude higher than the measured mass. In order to have a theory consistent with the measured Higgs mass, extreme fine-tuning is necessary to cancel out the divergences. While not technically forbidden or inconsistent, corrections over many orders of magnitude are worrisome.

Because of the current SM success, it is well motivated that a simple extension to the SM is a reasonable path to explain the aforementioned phenomena. These theories are known as Beyond the Standard Model (BSM) theories. As mentioned in the previous chapter, the Higgs is an area of the SM most recently discovered that does not have strong constraints on its decay modes. Thus, an exciting and natural class of BSM physics could reasonably come from modifications to the Higgs sector in the SM in the form of non-SM or "exotic" decays.

2.5.1 Exotic Higgs Decays

In addition to the limited experimental constraints on the Higgs decay channels, the Higgs has an extremely narrow width relative to its mass as a result of the decay processes being suppressed for various reasons.

Because of this, even a small coupling to new state can result in sizable branching ratios to new physics. These couplings can be on the order of 0.01 and still provide a BR to new physics of about 10%.

One of the simplest "exotic" Higgs decays is into a new light spin-zero particle, a. In many models where the a-boson mixes with the Higgs through Yukawa couplings, the dominant decay mode of the a is into a pair of b-quarks when the a mass is greater than the threshold for b-pair production. Thus, an appealing decay channel for new physics comes in the form of $H \to 2a \to 4b$.

These exotic decays are predicted by many BSM theories [9], including those with an extended Higgs sector (such as the Next-to-Minimal Supersymmetric Standard Model, NMSSM [10–14]), several models of dark matter [15–19], models with a first-order electroweak phase transition [20, 21], and theories with neutral naturalness [22–24].

A new spin-zero particle could be the mediator for dark-matter interactions with SM particles. This possibility is well motivated by the observations of the Fermi Large Area Telescope (FermiLAT) of a gamma ray excess originating from the Galactic Center (GC) [25–33]. The FermiLAT collaboration has recently confirmed this excess using all the data currently available and the most recent interstellar emission models for γ-rays [34] and has observed another similar excess in Andromeda [35]. This excess may be the first indirect observation of dark matter to date and is consistent with the annihilation through a light spin-zero particle with Yukawa-like couplings to SM fermions [18, 19, 25–33, 36–39]. Other astrophysical sources, such as millisecond pulsars, have recently been proposed as explanations [40, 41] and the process responsible for this excess remains an open question. However, the annihilation of DM (X) through the channels $XX \to a \to 2b$ and $XX \to 2a \to 4b$ is expected in general grounds and should be experimentally tested. Models explaining the GC excess via a light spin-zero particle are largely unconstrained by existing collider and direct DM searches.

A new light spin-zero particle has several other theoretical motivations. This particle could make the electroweak phase transition first-order, allowing for electroweak baryogenesis as an explanation of the observed baryon-antibaryon asymmetry of the universe [20, 21, 42–45]. Such a new particle can also be related to the naturalness problem of the Higgs boson mass. In various Beyond the Standard Model (BSM) scenarios, the quantum divergences in the Higgs boson mass are canceled by new particles related to the SM through new symmetries. In some cases, these new particles hadronize into spin-zero composite particles that have the same properties as a [23, 24]. These scenarios are experimentally untested and this search measurements can provide the some of the first constraints on a solution to the naturalness problem that does not contain new colored particles, *neutral naturalness*.

The search for signatures of Higgs boson decays to a pair of spin-zero a particles at the LHC is one of the

simplest and best-motivated possibilities for "exotic" decays [9]. In many scenarios where the Higgs boson decays to a pair of new spin-zero particles a, the dominant decay mode is to a pair of b-quarks ($a \to bb$) for $m_{\mathrm{a}} > 10$ GeV, above the b-pair production threshold. Despite the experimental challenges associated with such a $4b$ signature ($H \to 4b$), this is expected to be one of the most sensitive decay modes [9, 46–48] and thus the focus of this thesis.

Chapter 3

The ATLAS Experiment at the LHC

The Large Hadron Collider (LHC) is one of the world's leading particle colliders designed to collide protons at a center-of-mass energy of 14 TeV. Along the accelerator lies multiple interaction points with detectors built around them to analyze various physics processes. The ATLAS experiment utilizes one of these detectors (aptly named the ATLAS detector) to probe fundamental physics. The work described in this thesis uses data produced by the LHC and recorded by the ATLAS detector and the following chapter will detail the main aspects of both.

3.1 The Large Hadron Collider

The LHC [49] is a circular superconducting particle accelerator installed in a 27 km underground tunnel designed to collide protons. The accelerator ring houses four major detectors at various interaction points: ATLAS, CMS, LHCb, and ALICE. ATLAS and CMS probe a similar range of physics from the proton-proton collisions, while LHCb and ALICE focus on other specializations (b-physics and heavy ion collisions, respectively). For the ATLAS data considered in this thesis, the LHC was operating at a center-of-mass energy $\sqrt{s} = 13$ TeV.

3.2 The ATLAS Detector

The ATLAS detector is one of the main apparatuses used for detecting and analyzing the final decay products of pp collisions. It covers nearly the entire solid angle around the collision point and is comprised of several subdetectors that each have a dedicated role in processing specific expected decay products. The closest system to the beamline is the inner detector (ID), which is used to capture information about charged particles and their paths. Surrounding the ID is an electromagnetic calorimeter designed specifically to record signatures from electrons and photons. Another calorimeter designed to identify and measure energy of hadrons is also employed. Furthest out from the interaction point is the muon spectrometer (MS), which is dedicated to measuring muons. Before describing these subsystems in detail, a note about the ATLAS

coordinate system should be discussed.

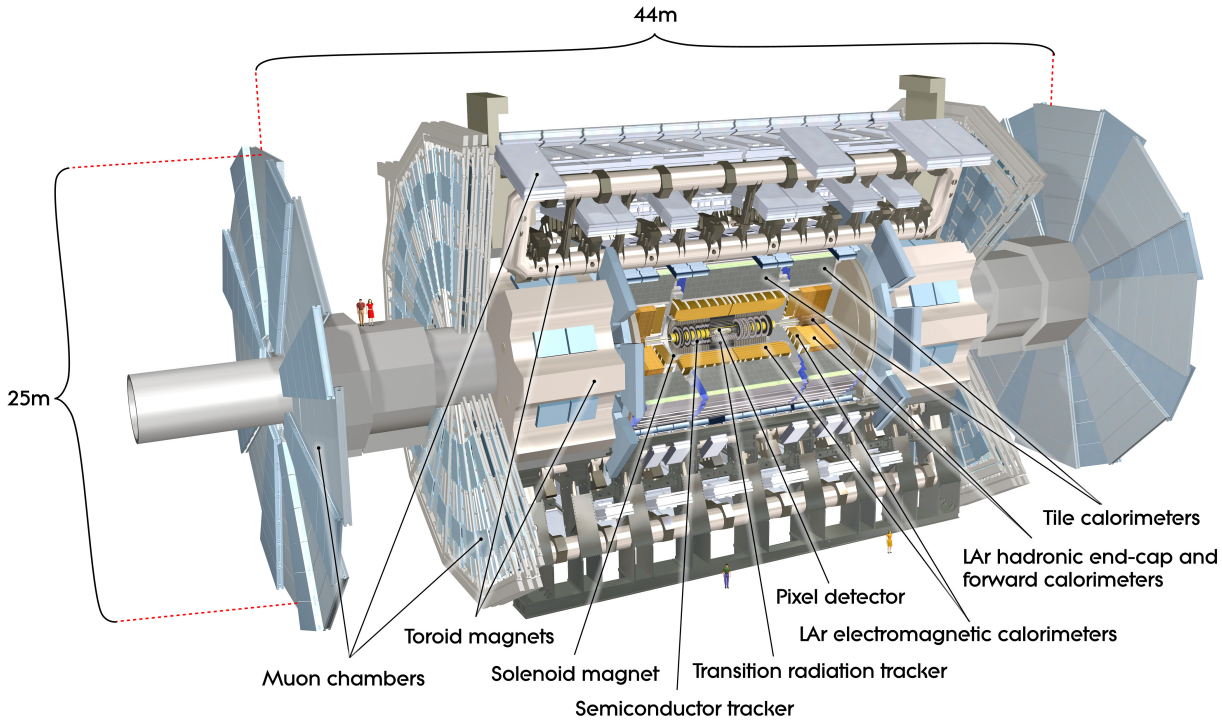

Figure 3.1: View of the full ATLAS detector [49].

3.3 ATLAS Coordinate System

The ATLAS coordinate system is a cartesian right-handed system with the origin and the interaction point in the center of the detector. The positive directions for x, y, and z correspond to pointing towards the center of the ring, upwards to the sky, and counterclockwise along the ring, respectively. Often the angles with respect to the beam axis are more useful. The azimuthal angle ϕ is measured around the beam axis, while the polar angle θ is measured with respect to the z-axis.

Since momentum in the z-direction cannot be determined, along with the fact that energy/momentum must be conserved in the x- and y- directions, it is useful to use quantities that are boost invariant along the z-axis or their transverse components. In particular, the pseudorapidity is used:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{3.1}$$

along with p_T and E_T, the transverse components of momentum and energy. Another useful quantity

commonly used is the angular distance between points is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.2}$$

One important caveat is that the pseudorapidity is only boost invariant along the z-axis if the particle is massless. But relative to the scales of energy/momentum, this approximation for massive decay products is valid in many cases.

With the coordinate system and useful parameterizations defined, the subsystems will be described in the following sections.

3.4 Inner Detector

The innermost layer of ATLAS is the inner detector (ID), which is used to reconstuct charged particles through energy/momentum and vertex measurements. The ID is immersed in a strong magnetic field and is comprised of three components (in order of distance from the beam axis): pixel, semiconductor tracker, and transition radiation tracker. The active parts of the ID record the particle's trajectory as it moves through the detector. These trajectories get reconstructed as objects referred to as tracks. Together with fast-response electronics, resilience to radiation, and the fact that there is little material between the beam and detector, the ID provides accurate measurements of charged particles covering a range of $|\eta| < 2.5$.

3.4.1 Pixel Detector

The pixel detector is the inner most component of the ID and is comprised of roughly 80.4 million readout channels, allowing for high granularity measurements within $|\eta| < 2.5$. The resolution achieved is 10 μm in the $R - \phi$ plane and 115 μm in the z-direction. This high precision is primarily used for vertex reconstruction. An additional pixel layer called the the Insertable B-Layer (IBL) was added as an upgrade to improve performance of the pixel detector. Due to its proximity to the beamline, the IBL allows for precision reconstruction of secondary vertices, an important object used for identifying b-quark final states, which is a critical component to the work described in this thesis.

3.4.2 Semiconductor Tracker

The next section further out from the interaction point is the semiconductor tracker (SCT). It is composed of a barrel that is lined with silicon strips parallel to the beam axis and endcaps with radial strips for precise

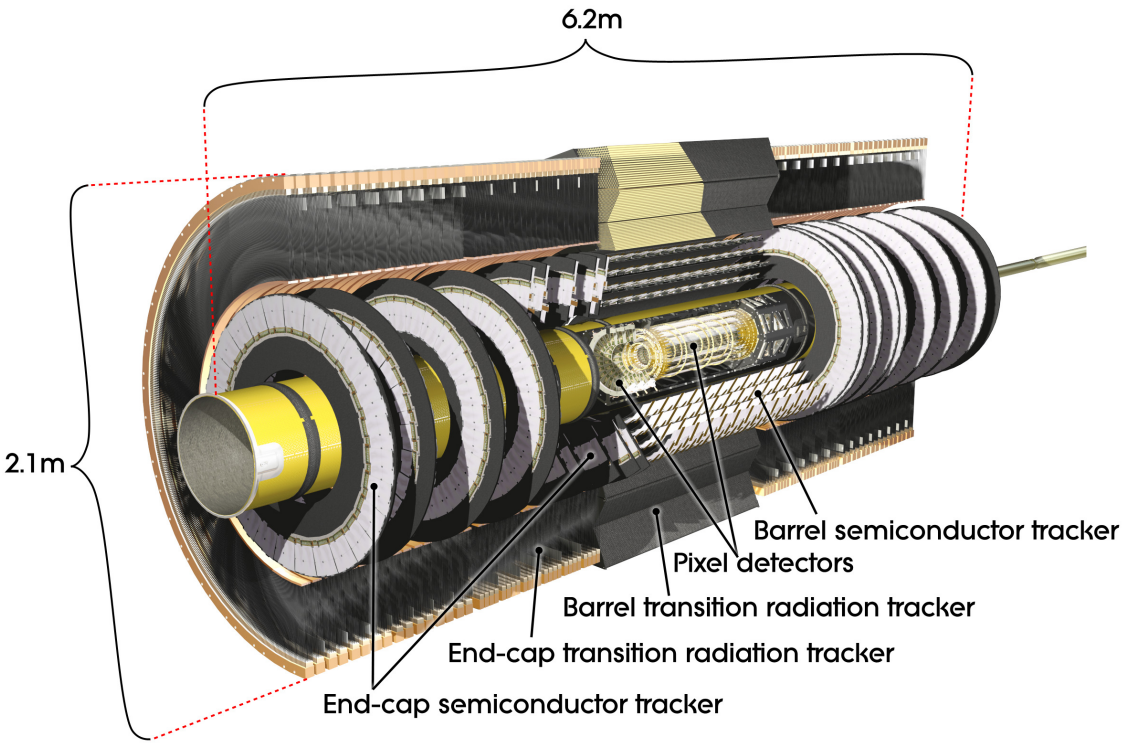

Figure 3.2: Cut-away view of the ATLAS Inner Detector [49].

$R - \phi$ measurements covering $|\eta| < 2.5$. There is also an additional layer of strips with a 40 mrad angle between the first layer to determine positioning of the tracks within a strip longitudinally using the crossing point of the strips between the different layers. The SCT mainly contributes to momentum reconstruction.

3.4.3 Transition Radiation Tracker

Surrounding the SCT is the transition radiation tracker (TRT). It is comprised of 4mm diameter straw tubes that are also placed parallel to the beamline as well as radially in the endcaps, but only cover a range of $\eta < 2.0$. The tubes contain a gaseous mixture of carbon dioxide and xenon, along with a gold-plated tungsten wire. Charged particles ionize the gas, and the ionization charges drift to the wire for detection. The time of arrival of the charges provide spatial information. The material between the tubes also produces radiation as electrons pass through. Thus, the TRT can aid in the identification of electrons.

3.5 Electromagnetic and Hadronic Calorimeters

The ATLAS calorimeter system is divided into two major components: electromagnetic (EM) and hadronic. The calorimeters contain a barrel and endcap and together cover up to $\eta < 4.9$. The EM calorimeter is comprised of an accordion structure with liquid argon active material and lead absorber plates. They are designed to stop most particles except muons and neutrinos. The EM calorimeter focuses mostly on absorbing and measuring energy from electrons and photons, while the hadronic calorimeter focuses on hadronic activity. The EM calorimeter has finer segmentation in the η region that the ID covers in order to have more precise measurements of photon and electron showers. The hadronic calorimeter and endcaps are not as granular, but are still sufficient for capturing hadronic activity (i.e. jet kinematics).

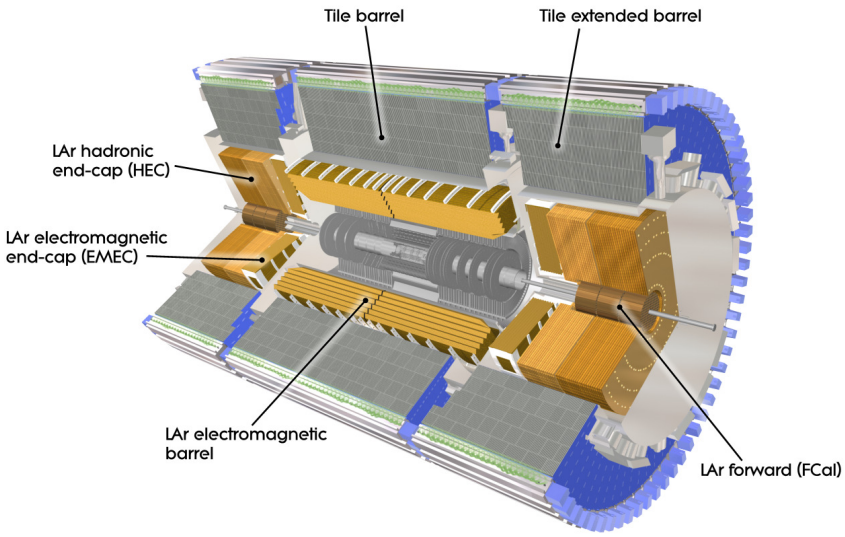

Figure 3.3: Schematic view of the ATLAS calorimeter system [49].

3.6 Muon Spectrometer

The Muon Spectrometer (MS) is the outer most section of the ATLAS detector comprised of a barrel and endcap region. As the name suggests, the MS is designed to identify and measure properties of muons. It is made up of four major detector technologies: monitor drift tubes/chambers (MDT), cathode strip chambers (CSC), resistive plate chambers (RPC), and thin gap chambers (TGC).

3.6.1 MDT

The MDTs are similar to the TRTs described in the Inner detector. They are comprised of tubes filled with argon and carbon dioxide gas with a tungsten-rhenium wire producing an electric field. As muons pass through the tubes, the gas ionizes and attracts ionization charges to the wire to produce a signal. A total of 6 or 8 layers of tubes are used for precise track reconstruction in each chamber while reducing the background of fake muons. The MDTs cover a range up to $|\eta| < 2.7$.

3.6.2 CSC

CSCs are multiwire proportional chambers positioned at high $|\eta|$ to help combat the higher rate and background conditions. The chambers contain the same gas composition as the MDTs. One advantage of this technology is the lower drift time, allowing for faster readout rates. The CSCs cover a range of $2.0 < |\eta| < 2.7$.

3.6.3 RPC

RPC chambers are made up of two plates of resistive material filled with a mixture of $C_2H_2F_4$, $\mathrm{Iso}-C_4H_10$, and SF_6 gas. When charged particles fly through the gas, avalanches are produced that get collected by two sets of orthogonal strips which provide the ϕ measurement in addition to η. RPCs have a very low drift time compared to both the CSCs/MDTs and are thus used in fast decision making for storing collision data, referred to as triggering (described in more detail in 3.7).

3.6.4 TGC

The TGCs are also multiwire proportional chambers designed for fast readout times like the RPCs, but capable of withstanding the higher rates and background conditions like the CSCs. The fast signal detection is due to the anode-cathode distance being smaller than the adjacent anode wire distance. The TGCs exist in the endcaps and cover a range of $1.05 < |\eta| < 2.7$. The TGCs also provide a measurement for ϕ in addition to η.

3.7 Trigger System

The bunch crossing rate is about 30 MHz when the beams are operational. Due to the technical limitations of the ATLAS detector and the computing resources necessary to record all those collisions, a triggering system

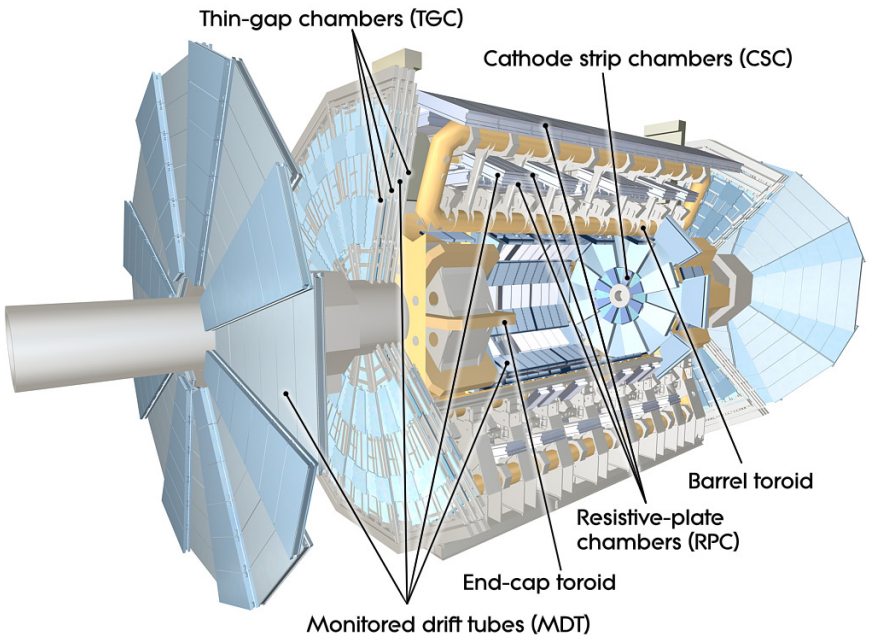

Figure 3.4: Schematic view of the ATLAS muon spectrometer [49].

was designed to make fast real-time decisions that save events with physics relevance. Most triggers are a set of conditions looking for candidates of quickly reconstructed physics objects like electrons, muons, jets, or even event-level properties like the total missing transverse energy (MET). The trigger system drastically reduces rate of input data to about 1 kHz.

The trigger system is comprised of 3 levels, each with increasing detailed information about the objects reconstruction. The first level (L1) is a hardware based system using coarse calorimeter and muon information. The second (L2) and third (event filter, EF) are software based systems that use information from all of the subdetectors. The last two levels combined are collectively referred to as the High Level Trigger (HLT).

The L1 trigger system utilizes fast, custom electronics to quickly identify potential physics objects by finding regions-of-interest (ROI). ROIs are computed in terms of η (or θ) and ϕ, then feed the raw data of the event to the next trigger level. The L1 trigger reduces the event rate down to < 100 kHz.

At the HLT, the L2 system uses the ROI inputs combined with tracker information to build candidate physics objects (electrons, muons, etc) and compute their position and energy. A selection is then applied to these objects to reduce the rate even further. After this selection, the physics objects are reconstructed using the full techniques that are used for analysis (sans detailed calibrations). Another selection is applied to these objects and the finally reduces the event rate to 1 kHz.

The selection applied at the HLT trigger is defined by a trigger menu, which is often updated and refined as the experiment progresses to accommodate all the physics analyses ATLAS is interested in. Some of the triggers are very inclusive (e.g. single objects with low p_T cuts). In some cases, the event rate is still too high to record everything. In order to further reduce the rate, triggers may have a prescale value. The prescale value is defined as the rate of acceptance for a given trigger, i.e. one out of every P events triggered will be saved, reducing the value by a factor of 1/P.

The nomenclature for a given trigger is formatted as:

$$\text{[LEVEL]\_[N][TYPE(S)][THRESHOLD]\_[ISOLATION]\_[QUALITY]} \tag{3.3}$$

with each component defined as: the trigger level used, the multiplicity of the physics object, the type of physics object (electron, muon, jet, MET), the p_T or energy threshold, requirement of isolation from other objects, and severity on final algorithm selection.

3.8 New Small Wheel

The LHC has plans to upgrade the accelerator complex to increase the rate of collisions and deliver more data in the coming years. ATLAS will also need to upgrade the various subsystems to accommodate the higher event rate and properly record high quality data. The higher event rate leads to higher background rates and places more demand on the processing required of the trigger systems. One important upgrade to the system is the New Small Wheel (NSW) whose main goal is to improve the muon triggers in the endcap region $1.3 < |\eta| < 2.4$.

3.8.1 Oveview of the NSW trigger

The current endcap muon trigger uses the TGCs in the middle layer of the detector, the big wheels (BW) [50]. The strategy uses coincidences of hits pointing to the interaction point (IP) in multiple layers of TGCs. Large trigger rates have been observed in the TGCs mainly due to secondary particles originating in the material near the beampipe in the region of the endcap toroid, which is location next to the TGCs.

The proposed upgrade is to replace the innermost layer of muon detectors, the small wheel (SW), which is located on the other side of the endcap toroid, with new detectors and electronics, the NSW [51]. The NSW is planned to provide an additional segment to the trigger logic to be combined with the TGC information (i.e. more high quality hit coincidences) and thus reject most of the background from secondary particles. The NSW segments should project back in θ to the IP with a precision of about 1 mrad.

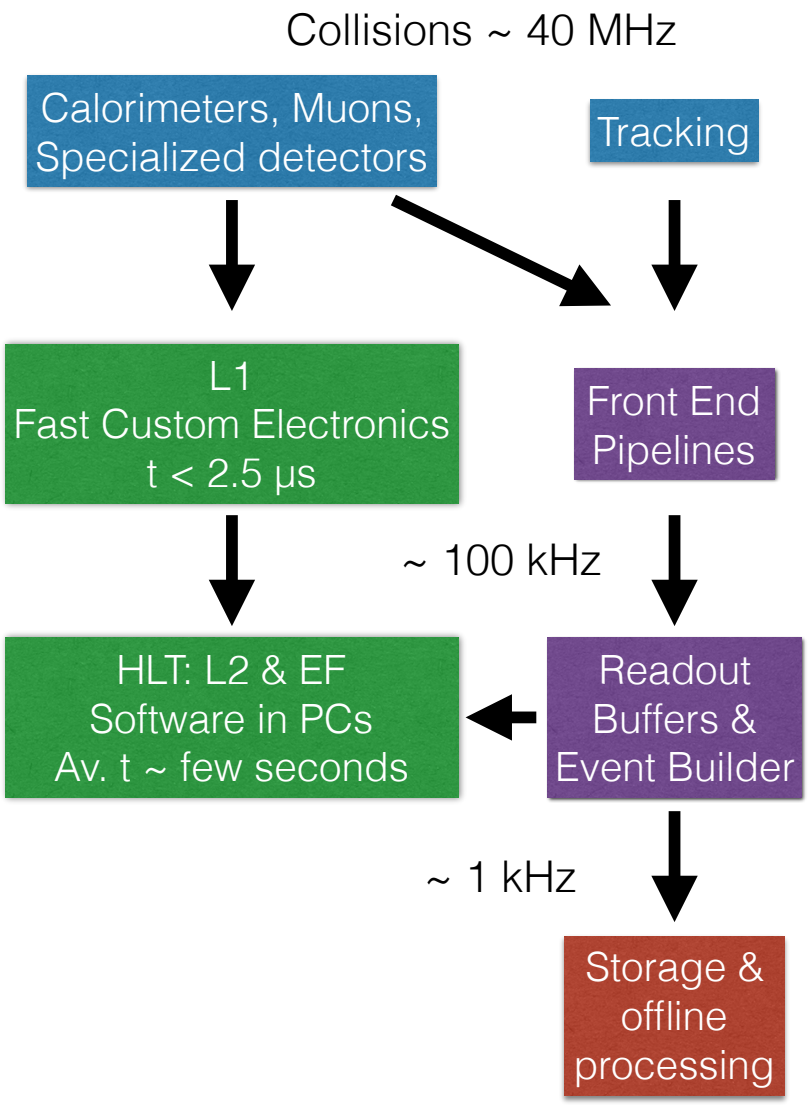

Figure 3.5: Schema of the ATLAS trigger system.

A key piece of information needed for the trigger is the pointing to the IP, $\Delta\theta$, which identifies particles originating from the collisions and discriminates them from particles originating from other sources. The trigger requires that the segment satisfies $|\Delta\theta| < 15$ mrad and also provides information about the position of the segment in η and ϕ so that the positions can be matched with the TGC information.

Figure 3.6 (left) shows examples of tracks accepted by the current trigger and indicates how the additional information will reject fakes. The existing trigger, based on the BW alone, would accept all three tracks indicated. With the addition of the NSW, only track 'A', the desired track, which is confirmed by both the Big Wheel and the NSW, will be accepted. Track 'B' will be rejected because the NSW does not find a

track that comes from the interaction point (IP) that matches the Big Wheel candidate. Track 'C' will be rejected because the NSW track does not point to the IP. The NSW logic restricts $\Delta\theta$ to a value consistent with a track that originates in the IP. Figure 3.6 (right) shows the expected reduction the triggers in the region $1.3 < |\eta| < 2.4$ with the NSW.

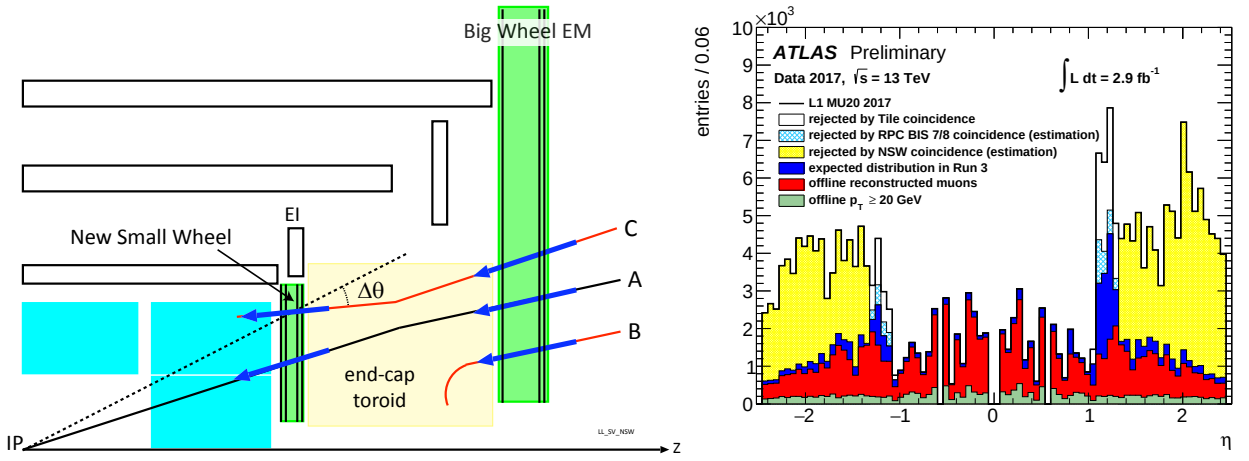

Figure 3.6: (left) Schematic of the muon endcap trigger indicating examples of which triggers in the current system would be rejected by the NSW. The NSW logic restricts $\Delta\theta$ to a value consistent with the track to have originated from the IP. (right) Expected reduction in triggers from the NSW in the region of $1.3 < |\eta| < 2.4$.

The NSW is segmented in ϕ as 16 trapezoidal sectors. Each sector is comprised of two detector technologies, resistive strip MicroMegas (MM) and small-strip Thin Gap chambers (sTGC). The geometry for a given sector is comprised of eight planes of MM detectors sandwiched between four planes of sTGC on each side along the z-axis. The readout scheme is comprised of a set of strips for the MM, and a set of strips, wires, and low granularity pads for the sTGC. Both technologies are planned to be used for tracking and for the trigger. For the sTGC, the wires and strips for tracking, while the strips and pads are used in the trigger. For the MM, all strips are used for tracking, while only the strip with the first hit in time above threshold for each 64 channels is used in the trigger. Four out of the eight MM planes have strips at a slight angle (referred to as the stereo angle) of $\pm$ 1.5 degrees in order to compute the azimuthal direction.

The MM and sTGC trigger algorithms run independently and the trigger segments are combined into a single set of candidates that are sent to the sector logic (SL) which then combines the information from the NSW with the TGCs in the BW. The NSW segments are computed with FPGAs on trigger processor boards which receive the trigger hit information from upstream electronics. A major challenge of the system is to be able to provide the segment candidates in the short latency of the system, which is 43 LHC bunch crossings or 1075 ns.

The following sections detail the work contributed to the NSW project, which is focused on the imple-

mentation of the MM trigger algorithm into the ATLAS simulation software.

3.8.2 The MM Trigger Algorithm

The MM trigger algorithm runs on the trigger processor FPGAs. The implementation of its simulation is in the ATLAS Athena software framework. Athena is the official framework for managing almost all ATLAS production workflows: event generation, simulation, reconstruction, processing formats for analysis, and the High Level Trigger. The goal of the work is to study the performance of the trigger and provide studies with the trigger integrated in the full detector simulation to use in realistic simulations for analysis once the NSW detector is operational. This contribution to the NSW project is the first implementation of the MM trigger algorithm in Athena.

The MM trigger algorithm receives inputs of MM trigger hits, referred to as Address in Real Time (ART) hits. As a muon passes through a MM plane, ionization charges get produced and drift to the readout strips. Multiple readout strips activate and produce hits. Each front end chip located on the MM detectors, called the VMM, takes hits from up to 64 MM strips and the address of the first strip in time to cross the threshold is the ART hit.

The data from the MM planes depends on the geometry. As briefly mentioned in the previous section, four planes have horizontal strips, referred to as X planes, while four planes have a slight stereo tilt to provide an azimuthal measurement, referred to as U and V planes, depending on the direction of the tilt (+1.5 degrees for U, -1.5 degrees for V). The eight MM planes are arranged in a XXUV-UVXX geometry along the z-axis.

3.8.2.1 Algorithm Description

The MM trigger algorithm begins by translating the addresses of the ART hits from each plane into slopes and associating them to predefined "slope-roads" using a look-up-table (LUT). A circular buffer is used to look for coincidences between the roads in the planes. Finally, the local and global slopes for track candidates are computed. The trigger algorithm can be outlined in four steps:

1. Translation of hardware addresses into equivalent track slopes fixed to the IP.
2. Determination of the presence of a multi-plane coincidence.
3. Parallel calculation of global θ (zenith of the track position at the entrance of the NSW) and local θ (direction, at the entrance of the NSW) angles with parallel strips and global average stereo strips, using the multi-plane coincidence.

4. Calculation of $\Delta\theta$ (the difference in θ between the direction of the segment extrapolated back to the interaction point and its direction when entering the detector region), global θ (referred to as θ), and ϕ (the azimuth defined with respect to the center (y) axis).

The first two items are performed by many *finders*, which don't consume significant amount of resources, but reduce significantly the throughput towards the second half of the algorithm. The latter two items are performed by *fitters* that consume most of the resources allocated to the algorithm, but only performed upon the presence of a solid track candidate.

3.8.2.1.1 Decoder

Incoming strip hit addresses are decoded into global slope values using by multiplying by a constant. A strip's stored slope value is defined as the orthogonal distance between a given strip and the beam line divided by the z distance from the IP to the plane the strip belongs to. It is precomputed taking into account a strip offset and a z position stored for each of the 8 planes and 16 radial segments of each sector.

3.8.2.1.2 Finder

The track finding algorithm is based on the LUT slope-roads coming from dividing the sector from bottom to top along η. Hit data from the decoder is collected and associated to a slope-road in the LUT. Details of the hit that can't be calculated from the road itself are also stored. The buffer mentioned in section 3.8.2.1 can hold the roads associated to a hit for each plane, with the hits expiring after the hit integration time. The buffer checks all slope-roads associated to the hits once per bunch crossing and determines if a coincidence threshold on the minimum number of planes with hits along the same slope-road has been met. Coincidence also requires the oldest hit of the track to be expiring (i.e. older than one bunch crossing). The strip number and slope for each hit are calculated and passed to the track fitting algorithm.

3.8.2.1.3 Fitter

In the fit, individual hit slopes in a slope-road are used to calculate global slopes associated with each plane type, which are averages (e.g. M_X^g for the average slope of horizontal planes). These in turn are used to calculate the three composite slopes: slopes associated with the horizontal (m_x) and vertical coordinates (m_y) and the local slope of hits in the horizontal planes (M_X^l), all of which are shown in:

$$m_x = \frac{1}{2}\cot\theta_{st}\left(M_U^g - M_V^g\right),\ m_y = M_X^g,\ M_X^l = \frac{\bar{z}}{\sum_i z_i^2 - 1/n\left(\sum_i z_i\right)^2}\sum_i y_i\left(\frac{z_i}{\bar{z}} - 1\right) \tag{3.4}$$

where θ_{st} is the stereo angle, the sums are over relevant planes, $\bar{z}$ is the average position in z of the horizontal planes, and y_i and z_i in the local slope expression refer to the y and z coordinates of hits in X planes.

From these composite slopes, the familiar expressions for the fit quantities θ, ϕ, and $\Delta\theta$ (approximated in 3.5) can be calculated as described in [52]:

$$\theta = \arctan\left(\sqrt{m_x^2 + m_y^2}\right), \ \phi = \arctan\left(\frac{m_x}{m_y}\right), \ \Delta\theta \approx \frac{M_X^l - M_X^g}{1 + M_X^l M_X^g} \tag{3.5}$$

These variables give an ROI that can be passed as a trigger signal.

Figure 3.7 shows the algorithmic flow of the finder and fitter with smaller functional units. Each of the steps, including the steps described for the finder in 3.8.2.1.2, shown in Figure 3.7 have been implemented as follows in both the hardware and in the corresponding software simulation.

(A) Incoming strip hit addresses are converted to global slope values.

(B) Hit slope values are stored in a circular buffer defined as (N slope-roads) $\times$ (8 planes) $\times$ (T), where T is the cyclical buffer depth and corresponds to the number of bunch crossings over which coincidences between planes are allowed. Candidates are identified when minimum hit threshold is met. Hits are kept in the buffer for a fixed number of bunch crossings.

(C) Each slope-road of the buffer is checked once per bunch crossing to determine if a coincidence threshold has been met.

(D) Slope-road contents containing the track candidate are read and cleared from the buffer and relevant track components are forwarded for processing.

Once a candidate track is identified, the following steps (E-I) to compute the ROI are completed in parallel:

(E) A local slope is calculated using a least squares fit of available horizontal-strip hits in the proposed track using the third expression in equation 3.4.

(F) A global horizontal hit slope, which is anchored to the IP, is calculated as the average of registered horizontal-strip hits in the proposed track candidate (M_X^{global}).

(G) A global stereo (U) hit slope, which is anchored to the IP, is calculated as the average of registered U hits in the proposed track candidate (M_U^{global}).

(H) A global stereo (V) hit slope, which is anchored to the IP, is calculated as the average of registered V hits in the proposed track candidate (M_V^{global}).

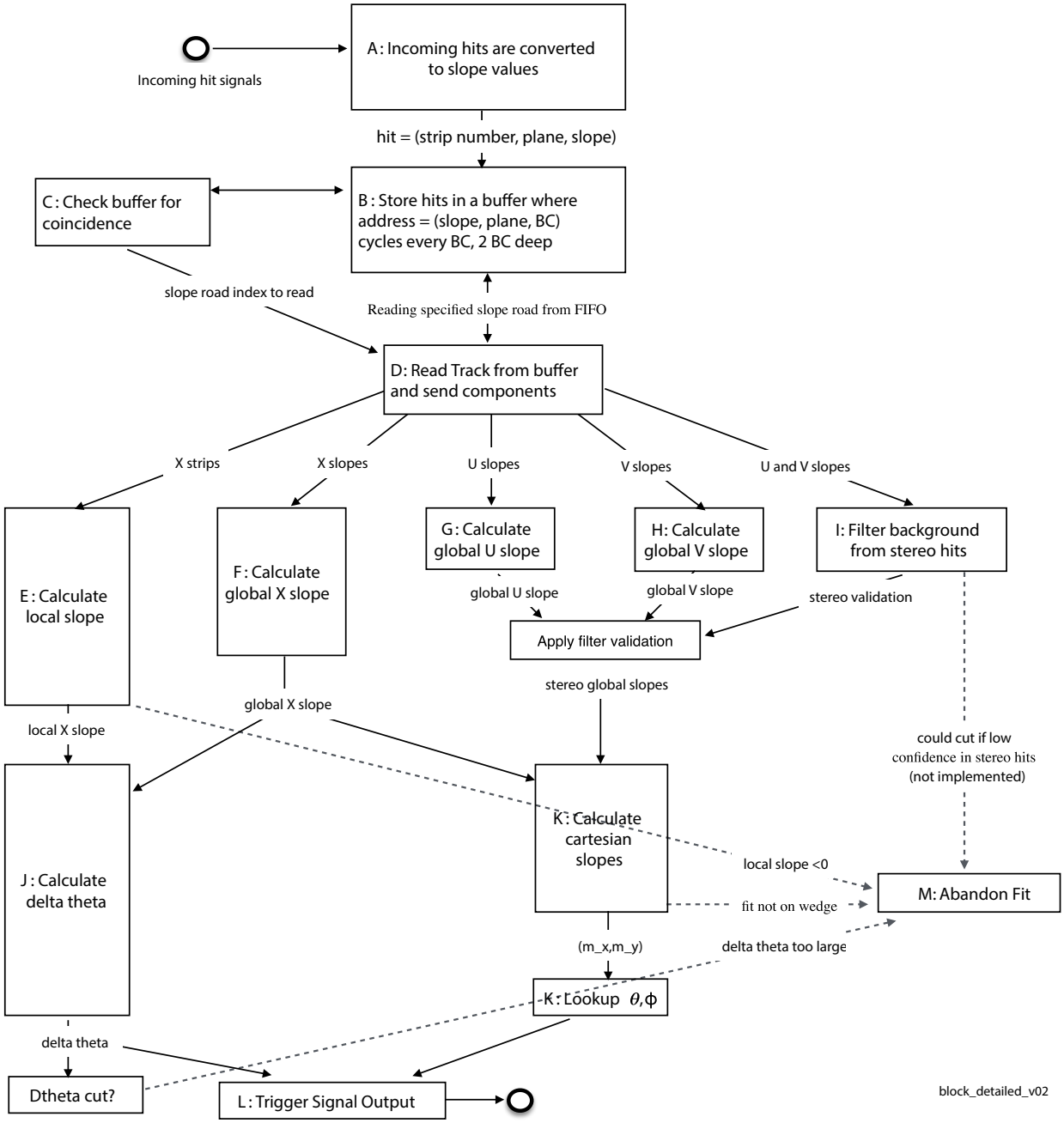

Figure 3.7: The block diagram is constructed with time flowing downward; therefore tasks on the same horizontal line are accomplished in parallel. Blocks correspond to operations comprising the algorithm, solid flow lines represent the flow of data, and light dotted lines represent fit abandonment signals, which can be triggered at multiple points throughout the algorithm. X in this diagram refers to horizontal strips, while U and V refer to the two sets of stereo strips (with a $+1.5°$ and $-1.5°$ stereo tilt respectively). Blocks after step D are approximately sized to represent their relative processing times.

(I) Stereo-strip background hits are further filtered from proposed tracks by judging how correlated two stereo-strip hits are with one another. In particular, strips with the same stereo tilt are compared (i.e comparing plane U_1 with U_2 and plane V_1 with V_2). If they are consistent between the two planes, the pair is kept, while otherwise it is discarded. If only one plane registers a hit on a strip of a given stereo tilt, it is kept.

(J) $\Delta\theta$ is calculated using previously fitted local and global horizontal slopes (M_X^{local} and M_X^{global}). This calculation is accomplished using a small ϕ angle approximation. In local coordinates, for which $\phi = 0$ at the middle of the sector.

(K) θ and ϕ are calculated using previously stereo and horizontal slopes. In particular, if hits exist for both stereo tilts, the cartesian position along the horizontal strip direction (m_x) is calculated using the two stereo hit slopes (U and V). If only one exists, the intersection point of the stereo strip and the horizontal strip is calculated and used. The other cartesian coordinate is given as $m_y = M_X^{\text{global}}$. The two cartesian coordinates are transformed into θ and ϕ using a look-up table. If the two cartesian coordinates do not correspond to a θ and ϕ in the sector (which can happen in cases with significant background contamination), the candidate is rejected.

(L) $\Delta\theta$, θ and ϕ are offered as a trigger signal.

3.8.2.2 Preliminary MM Trigger Performance

The performance of the algorithm is studied using single muon events generated with Athena full simulation. A geometry with two equal quadruplets each with two horizontal strips and two different-tilt stereo strips (XXUV-UVXX) was used. The coincidence requirement applied for the trigger is for 3/4 planes to have hits in X and 3/4 planes in UV. This configuration is referred to as 3X3UV.

The samples were generated using muons originating in the IP. Samples of muons with $p_{\text{T}} = 100$ GeV were used with a flat distribution in $1.3 < |\eta| < 2.4$ and ϕ in order to point to the NSW. No backgrounds are included in this simulation. Once the digitized hits are derived in the simulation, the first signal per channel is used for the trigger. The simulation includes the effect of the deadtime in the front end electronics and the coverage of the 64 strips associated to each VMM, ensuring that only one hit is used in the trigger for each grouping of 64 strips for each event. Once the algorithm was implemented successfully in Athena, the main goal of the studies was to study the expected performance.

The following variables are used to parameterize performance. The resolution is computed as the standard deviation of a gaussian fit to the distribution of the difference between the fit and true parameter. The true

definition is determined at the enetrans to the NSW region. Figure 3.8 shows the distributions for $\theta_{\text{fit}} - \theta_{\text{tru}}$, $\phi_{\text{fit}} - \phi_{\text{tru}}$, and $\Delta\theta_{\text{fit}} - \Delta\theta_{\text{tru}}$ and the fit results for the mean and standard deviation. Figure 3.9 shows the results of the fit for the resolution in bins of p_{T}. These results show that the resolution is within the specifications required for the trigger, though further studies are needed to improve the realism of the simulation and include other effects.

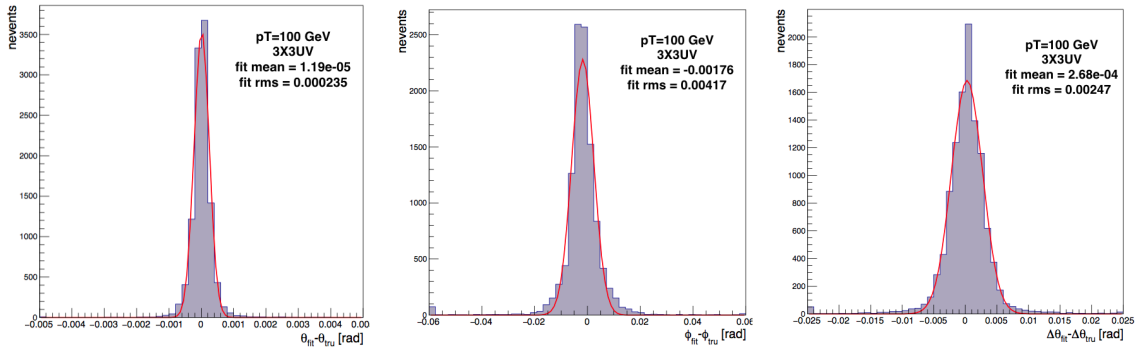

Figure 3.8: Distribution of (top) reconstructed θ minus true θ values, (middle) reconstructed ϕ minus true ϕ values, and (bottom) reconstructed $\Delta\theta$ minus true $\Delta\theta$ values of the track at the entrance of the NSW for muons of 200 GeV. The XXUV-UVXX configuration without background is used.

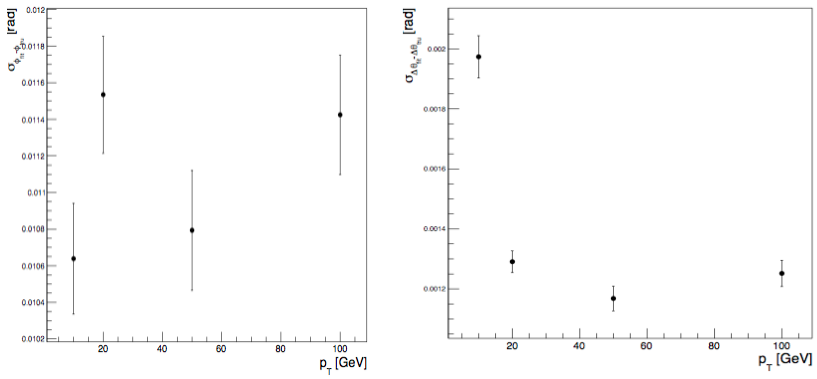

Figure 3.9: Distribution of the width of the (top) reconstructed ϕ minus true ϕ values and (bottom) reconstructed $\Delta\theta$ minus true $\Delta\theta$ values as a function of the p_{T} of the muon. The XXUV-UVXX configuration without background is used.

These studies correspond to the first implementation of the trigger algorithm in Athena. There is a campaign to update the realism of the inputs to the simulation, as well as update the finder portion of the algorithm to further reduce background impact by separating the finding of X and U/V coincidences. Future studies include studies fo the effects from incoherent backgrounds, which are expected to be large in the NSW region.

Chapter 4

Monte Carlo Simulation

Monte Carlo (MC) simulation is at the heart of nearly all ATLAS searches and measurements. In order to understand the data from the detector, detailed simulations of physical processes' (SM or otherwise) interactions with the detector are used to create samples of simulated data coming from various sources (signal or background). Comparisons between simulated and real data can then be made to infer what physical processes occurred.

As the name implies, MC simulation relies on Monte Carlo integration. This method approximates the expectation value of a function by taking the average value of a function after sampling many times using random number generators. Individual events can be simulated and produce an expected distribution.

The simulation from starting from physical process of interest caused by the pp collisions, all the way to the final detector signatures, is incredibly complex. An important feature that makes MC simulation possible is the fact that the full chain factorizes. Each major physical step can be simulated only having dependency on the previous step. This can practically be done using a Markov Chain model.

Each step is roughly broken up by the scale of the interaction. When the gluons and/or quarks collide and create the process of the interest (referred to as the hard-scatter), the momentum transfer is large and perturbation theory can be applied to describe the process. Any outgoing color charged particles will still be described by perturbation theory as they travel and radiate other color charged particles. Once the energy scale gets down to about 1 GeV, QCD perturbation theory does not hold and phenomenological models are used to describe the process of the quarks and gluons forming bound states and other low energy radiation. The final step is the simulation of the detector response as the final particles travel through the various sections of the simulated detector. Figure 4.1 shows a diagram of the process.

One very important property of MC simulation is that the *truth* information is saved throughout the entire process. Thus, MC samples have detailed information about how specific processes map to detector signals and how they would look in real detector data. This provides very useful information for understanding how different processes or sources contribute to the distributions in real data.

The following sections describe the MC steps in more detail.

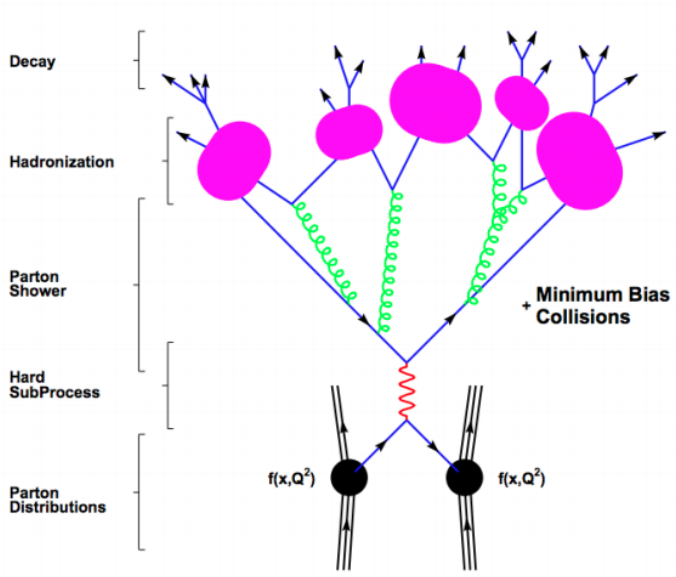

Figure 4.1: Diagram of the full MC simulation chain.

4.1 Factorization Theorem

The factorization theorem states that in hadronic collisions, the cross-section of a hard scattering process producing a final state X can be separated into short and long distance effects split by a factorization scale μ_F:

$$\begin{aligned}\sigma_{pp\to X} = \sum_{i,j} \int dx_1\, dx_2\, f_i(x,\mu_F^2)\, f_j(x,\mu_F^2) \\ \times\, \hat{\sigma}_{ij\to X}(p_1, p_2, \alpha_s(\mu_R^2), Q^2/\mu_F^2, Q^2/\mu_R^2)\end{aligned} \tag{4.1}$$

where the sum is over any partons that can cause X. The parton density function (PDF), $f_i(x,\mu_F^2)$, is a measured probability density function for a given parton i carrying the fraction of proton's momentum x. Since PDF's are independent of process, they are measured from many different experiments. The partonic cross-section $\hat{\sigma}_{ij\to X}$ can be expressed perturbatively as an expansion in α_s. This introduces a dependence on a renormalization scale μ_R which is typically set equal to the factorization scale. The computation of the partonic cross-section is what is referred to as the matrix element (ME) computation.

4.2 Matrix Element + Parton Showers

The partonic cross-section for producing X can generically be expressed as:

$$\hat{\sigma}_{ij} = \int d\mathcal{O} \underbrace{\sum_{k=0}^{\infty} \int d\Phi_{X+k}}_{\Sigma \text{ legs}} | \underbrace{\sum_{\ell=0}^{\infty} \mathcal{M}^{\ell}_{X+k}}_{\Sigma \text{ loops}} |^2 \tag{4.2}$$

where the sum k is over any additional hard emissions, the sum of ℓ is the sum of loop corrections, and $d\Phi_{X+k}$ is the phase space with k legs.

If k and ℓ are 0, this is referred to as the *leading order* (LO) computation. If $k > 0$, this computes the leading order cross-section for X with additional hard QCD radiation. If $\ell > 0$, this computes the next-to-leading order (NLO) cross-section. With each unit increment of ℓ, another "next" is added to the title, e.g. $\ell = 2$ is referred to as next-to-next-to-leading order (NNLO).

One important issue with the ME computation is that the k emissions must be high momentum (hard) and well-separated or else the integral does not converge. Since low momentum (soft) and collinear radiation is known to occur, these must be taken into account another way.

A separate simulation step referred to as parton showering (PS) is done to take into account the additional emissions as the partons travel. It is an iterative process that computes the probability of emission for a given parton by comparing the differential cross-sections of n to $n+1$ partons. The process runs until the scale of all the partons has gone to roughly 1 GeV.

One concern with simulating the PS independent of the ME is the potential overlap. If the ME calculation was done with additional legs and higher orders, there can be a double counting with the PS. There are many schemes to remedy this issue, but the general idea is to veto PS emissions that overlap with ME emissions and reweight the ME.

4.3 Hadronization

Once the transfer of momentum is on the order of about 1 GeV, perturbative QCD is no longer valid. The partons begin forming bound states and other models are necessary to describe the dynamics. In particular, the string fragmentation and cluster hadronization models are typically used to describe the parton behavior at low energy. These models have various tunable parameters to match data.

In the string fragmentation model [53, 54], the strong force from the gluons is represented as a string between quarks. As the quarks move farther away from each other, the string stretches, characterized by the

potential energy increasing. When the potential energy gets near hadron masses, the string breaks into two smaller strings by producing a quark/anti-quark pair. This continues until quark pairs with small strings are left.

In the clustering hadronization model [55, 56], the gluons after the parton shower are first split into quark/anti-quark pairs. The quarks are then clustered together to form colorless groups. These clusters are then allowed to decay into smaller clusters or hadrons, depending on their size.

4.4 Minimum Bias/Pileup

When collecting real data, the pp beam sends *bunches* of protons (roughly 10^{11} protons per bunch) that collide every 25 ns with about 20 actual interactions. This results in many low energy background scatterings that can be categorized as in-time pileup or out-of-time pileup. The former refers to additional interactions due to other pp collisions happening within the same bunch, while the latter refers to contributions from bunches adjacent in time. Since these interactions are low energy, similar phenomenological models are employed to simulate this behavior and create *minimum bias* events. Since minimum bias events are independent of the process of interest, they only need to be created and tuned to the conditions of the detector and beam the real data was taken under. They can then be overlayed on the simulated hard-scatter to provide a more accurate representation of the detector data.

4.5 ATLAS Simulation

The entire process described in the previous sections is practically produced with a set of various MC generators that each have their pros and cons depending on the physical process of interest. A brief summary can be found in 4.1 for reference.

Monte Carlo	**Matrix element**	**Parton shower**	**Hadronization**
Pythia6	LO	Parton shower	String model
Herwig	LO	Parton shower	Cluster model
Sherpa	LO	Pythia6	Pythia6
Alpgen	LO	Pythia6 or Herwig	Pythia6
MC@NLO	NLO	Herwig	Herwig
POWHEG-BOX	NLO	Pythia6 or Herwig	Pythia6 or Herwig
AcerMC	LO	Pythia6 or Herwig	Pythia6 or Herwig
MadGraph	LO	Pythia6	Pythia6

Table 4.1: Summary of the Monte Carlo generators used in the analysis.

The final output running the generators is a list of four vectors for stable particles after decay and hadronization. In order to create simulated data, the stable particles are processed through an extremely detailed simulation of the detector called Geant4 [57]. The particles leave behind energy deposits which get converted into electronic signals which can ultimately be reconstructed as physics objects.

Although MC simulations are very powerful, they are not perfect. Some detector effects and physics behaviors cannot be reliably simulated. In order to account for these mismodelings, calibrations are done to correct the MC behavior. By taking measurements in data samples that are very well-understood, discrepancies between MC and data can be reasonably alleviated.

Chapter 5

Overview of Statistical Procedure

This chapter introduces the statistical concepts used in the double b-tagger calibration and $H \to 2a \to 4b$ analysis.

5.1 The Likelihood Function

The first step in defining the statistical methodology is introducing a proper likelihood function, which gives the probability of an observation given a particular model. Data from the detector comes in the form of counts, and for the sake of simplicity, the example of a 1 bin (i) histogram will be used.

The expected number of events for bin i comes from MC simulation and can be expressed as:

$$E_i = \mu \cdot s_i + b_i \tag{5.1}$$

with s_i and b_i corresponding to the expected signal and background events from MC. Assuming the data coming from the detector follows a Poisson distribution, the likelihood can be expressed as:

$$L_i = \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \tag{5.2}$$

where n_i is the number of data counts in the bin. Various systematic and statistical errors can affect the model. These uncertainties are encoded in terms of *nuisance parameters* (NP), θ. Each NP is characterized by a probability density function $\rho(\theta)$. For most systematic errors, their PDF is assumed to be a Gaussian with a mean and width measured by prior experiments. For MC statistical errors, a Gamma distribution is assumed. Flat priors are sometimes used as a normalization parameter for specific backgrounds. This is typically done in analysis strategies that have dedicated regions, or sets of bins, designed to measure the background in a unique phase space where previous measurements or systematics may not have probed.

Combining all of these together gives a complete likelihood:

$$L = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \prod_{k}^{M} \rho(\theta_k) \tag{5.3}$$

where N is the total number of bins and M is the total number of systematics being considered. This likelihood can be maximized to fit both μ and all θ to give $L(\hat{\mu}, \hat{\theta})$.

After the fit, the NP can be affected in a few ways:

1. The fit NP central value can differ from the initial pre-fit central value, which is referred to as a *pull*.
2. The fit NP error can be reduced from the initial pre-fit error, indicating that assigned prior was too large and the measurement has *constrained* the NP.
3. If there is not enough statistical power in the measurement, neither pulls nor constraints may occur.

When plotting NP, they are typically centered at 0 with a pre-fit uncertainty normalized to 1 for ease of interpretation.

The likelihood can also be maximized with a fixed μ giving $L(\mu, \hat{\hat{\theta}})$. This negative log of this ratio (referred to as the profile likelihood ratio) gives the test statistic used for most searches at the LHC, including the $H \to 2a \to 4b$ search:

$$q_\mu = -2\ln\left(\frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}\right) \tag{5.4}$$

Before proceeding further, one important note should be made. This expression is useful when defining the likelihood for physics searches because the goal of the search is to identify new signals and perform hypothesis tests using the test statistic to draw conclusions about how nature behaves (described in the following section). For calibrations, the goal is different. Specifically, for the calibration of the double b-tagger does not look for a signal. The likelihood is used as a model that gets maximizes the likelihood by adjusting the MC. The parameter adjustments that best fit the MC to data are exactly the calibration scale factors used to correct the MC behavior to data. In this case, no "signal" or "background" is defined. A well-understood data sample is used to take a measurement to fix the MC that should model it. No likelihood ratio is necessary since no test statistic is being used for hypothesis testing. Precise definitions of the likelihood used in the calibration will be discussed in chapter 8.

5.2 Hypothesis Testing

In the case of the $H \to 2a \to 4b$ analysis, a search is being performed and requires the use of the test statistic q_μ for hypothesis testing. When aiming to discover a new process as the signal, the signal+background model is considered the alternate hypothesis (H_1) and the background-only model is considered the null hypothesis (H_0). In the event that there is no evidence for discovery, the null and alternative hypothesis definitions switch and limits are set to constrain the theory that predicts the signal.

From equation 5.2, it is clear that μ represents the signal strength. In other words, μ fixed at 0 would be considered the background-only hypothesis and $\mu > 0$ would be the signal+background hypothesis. The compatibility of the data under a given hypothesis μ' is given by the p-value:

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu')dq_\mu \tag{5.5}$$

where $f(q_\mu|\mu')$ is the PDF of q_μ for signal strength μ'. Estimations for $f(q_\mu|\mu')$ can be computed using Monte Carlo methods, but these are often computationally expensive. According to Wilk's theorem, if the fitted $\hat{\mu}$ is Gaussian distributed, the PDF of the test statistic asymptotically approaches:

$$\begin{aligned} f(q_\mu|\mu') = &\frac{1}{2\sqrt{q_\mu}}\frac{1}{\sqrt{2\pi}}\times \\ &\left[\exp\left(-\frac{1}{2}\left(\sqrt{q_\mu}+\frac{\mu-\mu'}{\sigma}\right)^2\right)+\exp\left(-\frac{1}{2}\left(\sqrt{q_\mu}-\frac{\mu-\mu'}{\sigma}\right)^2\right)\right]. \end{aligned} \tag{5.6}$$

where μ' and σ are the mean and standard deviation of $\hat{\mu}$.

In order to approximate σ, a pseudo dataset referred to as an Asimov dataset [58] is used. The Asimov dataset is constructed as a mock dataset defined to be equal to the predicted value (the sum of the background predictions). This provides a useful dataset to measure expected results and errors.

With a well-defined PDF for the test statistic, the p-value can be readily calculated for a given measured test statistic. However, one important issue occurs when the PDFs of both hypotheses are similar. A measured downward statistical fluctuation could result in the exclusion of both hypotheses. To remedy this, the CL_s method [59] was introduced as a ratio of probabilities:

$$CL_s = \frac{p_\mu}{1-p_0} \tag{5.7}$$

with p_μ and p_0 being the p-values of the compatibility of data under the signal+background and background-only hypothesis, respectively. The quantity $1-p_0$ is small when a downward background fluctuations occur

and bring the value of CL_s up. Thus, the CL_s method avoids excluding signals that a measurement or search is insensitive to. At the LHC, including the work of this thesis, the CL_s value is used instead of p_μ. Exclusion of a signal at the 95% confidence level is found when the $CL_s < 0.05$.

Chapter 6

Motivation for a Low Mass Double B-Tagger

After the discovery of the Higgs [2, 3], there has been on ongoing effort at the LHC to measure the properties of the newly discovered particle and search for new physics. As mentioned in section 2.5.1, the Higgs decays have not been fully measured and there is potential for new physics to couple to the Standard Model through the Higgs. A large effort exists to search for direct "exotic" Higgs decays that result from a class of models extending the Higgs sector [9]. One of the simplest extensions of the SM has the Higgs decaying into a new light spin-zero particle, a. In many models, the dominant decay mode of the a is into a pair of b-quarks when the a mass is greater than the threshold for b pair production. Thus, an appealing decay channel for new physics comes in the form of H $\rightarrow$ aa $\rightarrow$ bbbb. This chapter outlines the properties of the exotic Higgs decay signal and motivates that strategy for the work described in this dissertation.

6.1 H $\rightarrow$ aa $\rightarrow$ bbbb Modeling

The benchmark model considered for this thesis is an additional singlet (pseudo-)scalar a, with the following effective interactions:

$$\mathcal{L}^{\text{BSM}} \supset i y_b^a a \bar{b} \gamma_5 b + \frac{1}{2}\lambda_{aH} H a^2 + \frac{1}{2} m_a^2 a^2, \tag{6.1}$$

where m_a is the mass of this singlet scalar, and a Yukawa-type of coupling y_b^a controls the decay of the singlet a and the scalar quartic coupling λ_{aH}, which determines the partial width of the Higgs boson decaying into a singlet pair.

In the signal of interest, the energy of the Higgs boson is partitioned into multiple b-quarks. Thus, the final detector signatures for b-quarks, b-jets (described in the following chapter) have relatively low-p_T and may be overlapping. This is the key difference between this search and other searches with similar objects in the final state, with higher p_T b-jets well separated in the detector.

This multi b-jet final state that results from such signal models is challenging from the point of view of triggering. The most straightforward trigger strategy is to focus on associated production of the Higgs

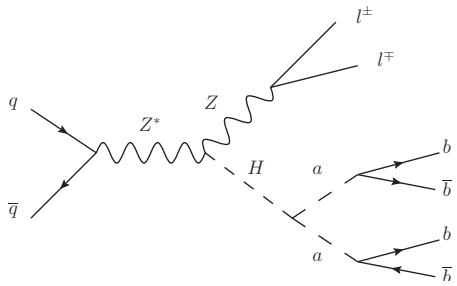

Figure 6.1: Representative leading-order Feynman diagrams for the Higgs boson production in association with a Z boson decaying to leptons and the exotic Higgs boson decay signal $H \to 2a \to 4b$ probed in this search.

boson with a Z boson, $pp \to ZH$, and use the leptons from the leptonic decays of the Z boson to trigger on the events. The experimental signatures of such a signal produced in the ggF and VBF Higgs boson production modes suffer from large QCD backgrounds, which are greatly reduced by the presence of leptons in the event.

A diagram for the signal of interest $H \to 2a \to 4b$ in the associated production mode ZH, is shown in Fig. 6.1. A generated MC signal sample is shown in the following sections to motivate the analysis strategy.

6.2 Signal sample generation

Signal samples of associated Higgs boson production with a Z boson, $qq \to ZH$ are generated with POWHEG-BOX v2 [60–63] using the CT10 PDF set [64]. The $pp \to ZH$ process includes $gg \to ZH$, which contributes over 10% of the total cross section. A Higgs boson mass of $m_H = 125$ GeV is assumed. The events are interfaced with the Pythia8 generator [65] that models the decay of the Higgs boson into a pair of new light scalars a with a mass set to $m_a = 20$ GeV. These new scalars a are then decayed into to a pair of b-quarks with BR$(a \to b\bar{b}) = 1$. Arbitrary branching ratios are obtained by reweighting samples using particle-level information. Events are subsequently showered and hadronized with Pythia8 using the AZNLO tune [66] for the modeling of the underlying event.

Events from minimum-bias interactions are simulated with the Pythia8 generator. They are overlaid on the simulated signal events according to the luminosity profile of the recorded data. The contributions from these pile-up interactions are modeled both within the same bunch crossing as the hard-scattering process and in neighboring bunch crossings. Finally, the generated samples are processed through a simulation [67] of the detector geometry and response using GEANT4 [57]. All samples are processed through the same reconstruction software as the data. Simulated events are corrected so that the object identification efficiencies, energy scales and energy resolutions match those determined from data control samples.

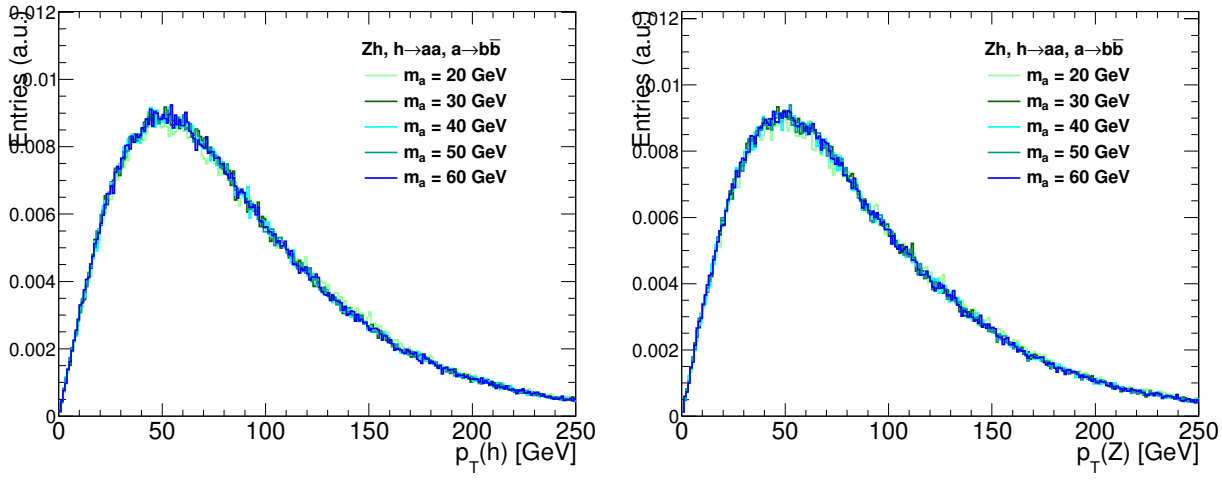

Figure 6.2: The p_{T} distribution of (left) the Higgs boson and (right) the Z boson for the ZH production mode using POWHEG-BOX.

The kinematics of the signal sample were explored to see the expected behavior of the decay products for the signal signature.

6.2.1 Signal kinematics

The studied kinematics of the signal samples are generated at truth-level, based on the information from the hard-interaction before showering. The signal samples are generated at NLO (Powheg-Box+ Pythia8) and agree very well with predictions generated at LO (Madgraph5+ Pythia6) at the parton and particle levels. For this comparison, MadGraph5 1.5.11 was used with the CTEQ6L1 PDF set [68] and Pythia6 for showering. The signal model is implemented using a modification of Madgraph's heft model to include an additional pseudo-scalar.

Figure (6.2) shows the truth-level Higgs boson p_{T} distribution for ZH production at $\sqrt{s} = 13$ TeV. The kinematics of Z and H are independent of the mass of the decay of the particle a.

Figures 6.3 and 6.4 show kinematic distributions for the leptons, the scalar a and the decay b-quarks for the ZH production mode. The separations for the heavier scalar a is smaller compared to the lighter one, since the p_{T} over mass is smaller. In the case of the separation between b-quarks originating from the decay of an a particle, for larger a masses, the quarks are more back-to-back, due to the smaller boost. Finally, it should be noted that these are relatively soft particles. For higher values of m_a, the decays tend to be more asymmetric and include lower momentum b-quarks.

These distributions highlight some key features of a search strategy for such a signal:

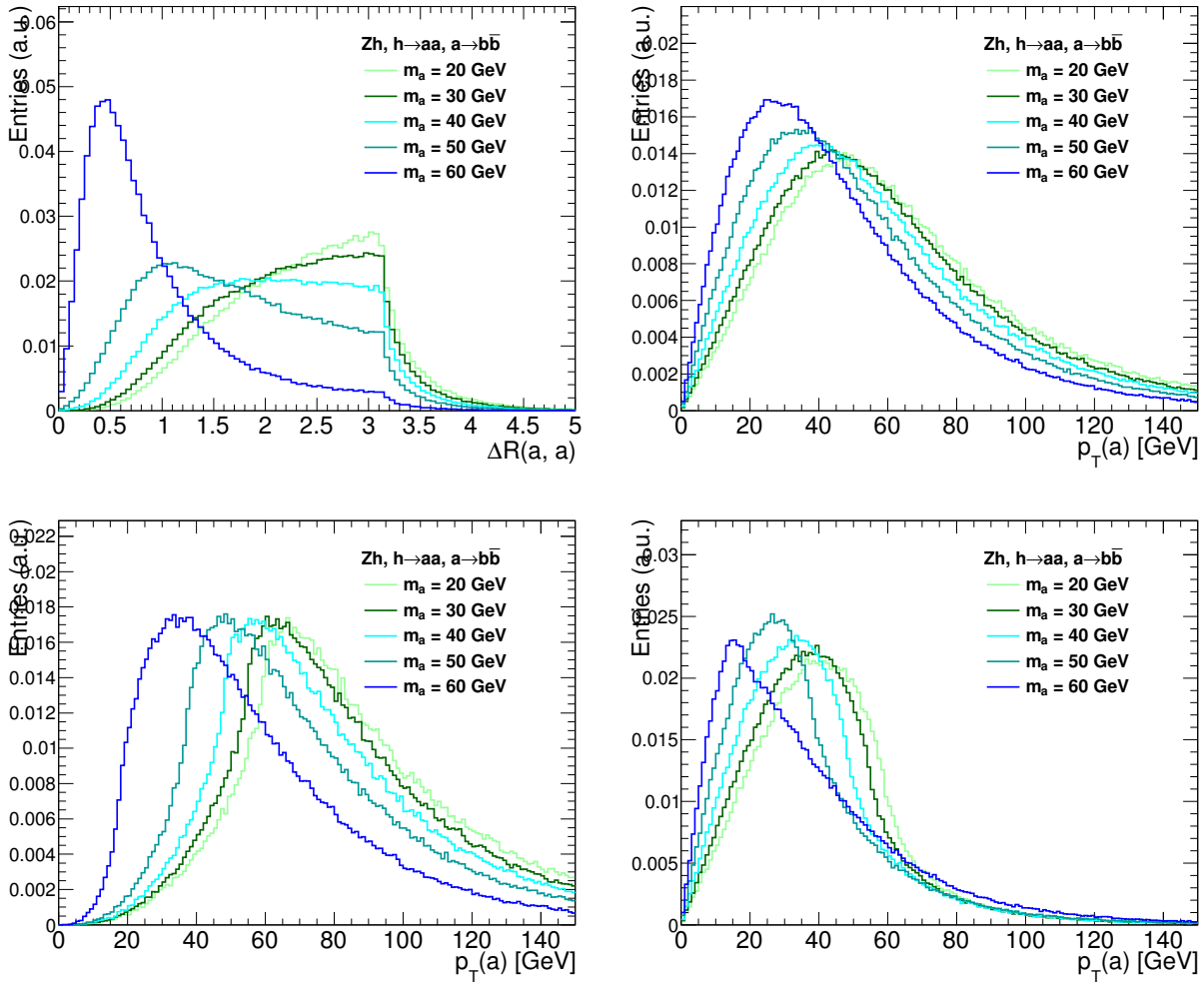

Figure 6.3: Kinematic distributions for a scalar particles from the $H \to 2a \to 4b$ signal for the ZH production mode. The figures show (top left) the separation between particles, (top right) the p_{T} distribution of both a bosons, (bottom left) the p_{T} distribution of the leading a boson, and (bottom right) the p_{T} distribution of the sub-leading a boson, for a range of masses m_a.

- For larger m_a, around half the Higgs boson mass, the decay products have typically higher p_{T} than for low masses, and are well separated. It is convenient for a **resolved analysis**, where the goal is to observe as many separate signal b-quarks as possible.

- For lower m_a, particularly below ~ 30 GeV, the decay products have relatively lower p_{T} and may be overlapping in the detector (as seen in the $\Delta R(bb)$ distribution). Furthermore, the p_{T} of the a particle is relatively higher in this case. It is appropriate for a dedicated **overlapping analysis**, where the goal is to reconstruct a low mass $b\bar{b}$ resonance as a single detector object.

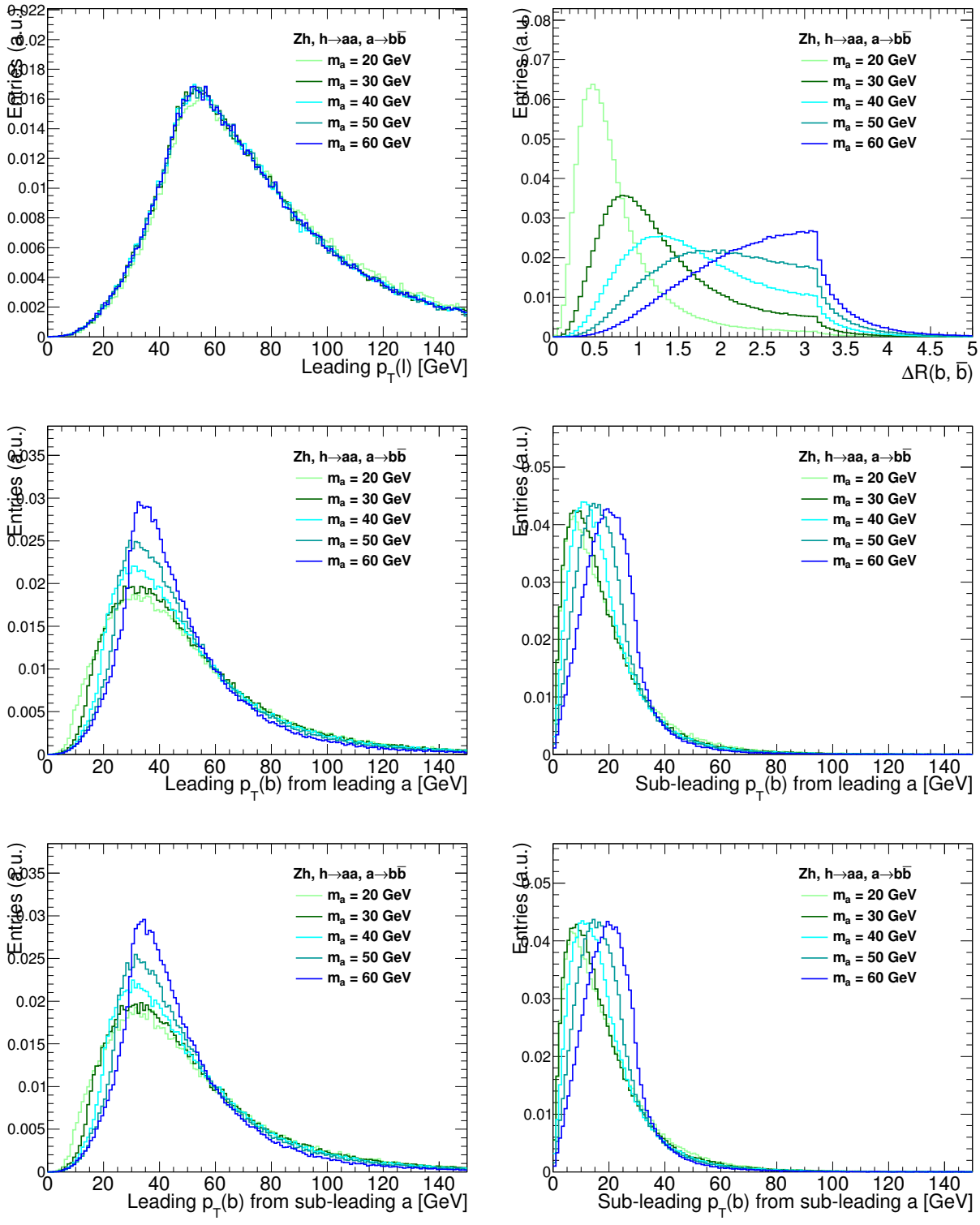

Figure 6.4: Kinematic distributions for the leading lepton coming from the Z boson and the b-partons from the $H \to 2a \to 4b$ signal for the ZH production mode. The separation between particles and the p_{T} distributions are shown for a range of masses m_a. Note that the separation is shown for pairs of particles from the same parent.

6.3 Previous results

A previous search for $H \to 2a \to 4b$ for m_a between 20 and 60 GeV was done using 2015 and 2016 data [69] using the resolved analysis strategy. Associated production with a W/Z production was investigated. The analysis focused on finding individual b-quark signatures in the detector (as well as leptons for the W/Z bosons). The limits set on on ZH production cross section times branching fraction for the considered mass range can be seen in Fig 6.5. At $m_a = 20$ GeV, a clear loss of sensitivity and degraded limits can be seen. This is understood to come from the effect of overlapping b-quarks leaving behind unresolvable detector objects. Thus, a new strategy is needed that requires a method to find signatures in the detector with two b-quarks in them. A technique for reconstructing an $a \to b\bar{b}$ resonance requires a dedicated tagger, which will be discussed chapter 8.

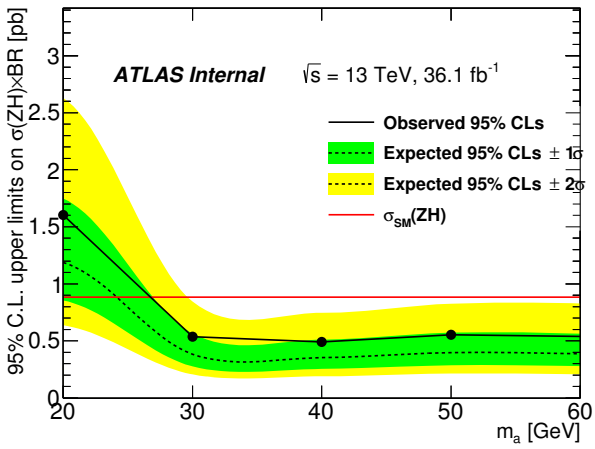

Figure 6.5: Previous limits for the range of m_a mass points 20-60 GeV on ZH production cross section times branching fraction.

Chapter 7

Object definitions

This chapter details the common objects in ATLAS used for the work in this thesis. The critical objects are jets, including the identification (tagging) of b-jets. Electrons and muons are also used mainly for triggering and event preselection. This section provides a summary of the reconstruction and identification of these objects. Other objects less commonly used in analyses, that are especially important for the work in this thesis, are introduced in the following chapter.

7.1 Electrons

Electron candidates [70–72] are reconstructed from energy deposits (clusters) in the electromagnetic calorimeter associated to reconstructed tracks in the inner detector. As electrons pass through the calorimeter, they interact with the detector and begin to shower photons and e^+e^- pairs. These will leave signals in the detector similar to jets, however, the shower shapes from electrons will be different compared to the charged decay products from a hadronic jet. The associated tracks to the EM clusters are also expected to be much cleaner, having higher numbers of pixel hits in the inner detector and lower impact parameters. There are two impact parameters that are important, transverse (d_0) and longitudinal (z_0). The transverse impact parameter is defined as the shortest distance between the track and the beamline, while the longitudinal impact parameter is defined as the distance in z between the primary vertex and the point on the track used to define d_0. Using shower shape variables and quality cuts on associated tracks, the amount of electrons faking jets can be reduced.

For the $H \to 2a \to 4b$ search, candidates are selected with $|\eta| < 2.47$, excluding the calorimeter transition region $1.37 < |\eta| < 1.52$. Electrons must pass the tight likelihood-based identification criterion `tightLH` [73], and are further required to have $|z_0 \sin\theta| < 0.5$ mm and $|\frac{d_0}{\sigma(d_0)}| < 5$, where the longitudinal and transverse impact parameters are computed with respect to the beam-line. The dilepton channel requires the presence of exactly two electrons, the leading one with $p_T > 27$ GeV and the sub-leading one with $p_T > 10$ GeV.

To further reduce the background from non-prompt electrons, conversions and hadrons, electron candi-

dates are required to be isolated and must pass the "Gradient" isolation working point [74]. This working point is defined to be a sliding scale isolation selection.

7.2 Muons

Muons are not expected to leave a strong energy signature in the calorimeters due to being a minimum ionizing particle. Instead, muons are reconstructed from tracking information from the inner detector and muon spectrometer. The most common type of muons used in analyses are *combined muons*, the collection that combines the tracking information from both the inner detector and muon spectrometer. Other muon collections exist with looser reconstruction criterion to improve acceptance if necessary.

Muon candidates are reconstructed from track segments in the various layers of the muon spectrometer, and matched with tracks from the inner detector (combined). The final muon candidates are refitted using the complete track information from both detector systems, and required to satisfy $p_T > 10$ GeV and $|\eta| < 2.5$. Muons are required to pass the medium quality requirements [75] and to satisfy the "Gradient" isolation working point. The absolute value of a muon's d_0 significance must be less than 3, and the value of $|z_0 \sin\theta|$ must be less than 0.5 mm.

For the $H \to 2a \to 4b$ search, a selection is applied to reduce backgrounds from muons from heavy flavor decays inside jets, muons are required to be separated by $\Delta R > 0.4$ from the nearest jet, removing the muon if the jet has at least three associated tracks, and removing the jet otherwise (this avoids an inefficiency for high-energy muons undergoing significant energy loss in the calorimeter).

7.3 Jets

As mentioned above, due to the nature of the strong force, a free quark is rarely found in nature. When a quark is produced, it radiates due to QCD processes and a cascade of additional gluons and quarks, called a *parton shower*, is produced. The color charge of the partons attracts other color charged partons and *hadronization* begins. As more radiation is produced, the energy scale decreases leading to stronger coupling strength, as described in Section 2. The partons form bound states and what results is a spray of hadrons that leave energy deposits in the calorimeters of the detector. The deposits are then reconstructed into *topoclusters* [76], which then get clustered with an algorithm into an object called a *jet*. The current ATLAS standard jet gets produced using the anti-k_T algorithm.

7.3.1 The Anti-k_T Algorithm

The anti-k_T algorithm [77] is a sequential clustering algorithm that organizes an input of four vectors and produces jets. These four vectors can come from topoclusters, tracks, or MC particles to form different types of jets (topo, track, and truth jets, respectively). From these four vectors the algorithm proceeds by computing a weighted distance between the constituents (d_{ij}) and between a constituent and the beam (d_{iB}):

$$d_{ij} = \min\left(\frac{1}{k_{ti}^2}, \frac{1}{k_{tj}^2}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{7.1}$$

$$d_{iB} = \frac{1}{k_{ti}^2}, \tag{7.2}$$

where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, k_{ti} is the transverse momentum of the i-th constituent, and R (referred to as the radius of the jet) is an input parameter of the algorithm that characterizes the size of the jet. As of now, radius $R = 0.4$ jets are the current standard in ATLAS.

The minimum of all the d_{ij} and d_{iB} computed is then found. If the minimum is a d_{ij}, then those two constituents are clustered by adding the four vectors together. However, if the minimum is a d_{iB}, the i-th constituent is considered a jet and the constituent is removed from the list of inputs. This process continues until there are no constituents left. The jet collections created from caloclusters are typically the ones used in analyses, but they cannot be used right after clustering for a number of reasons. A strong effort is in place to properly correct/calibrate the jets to have them ready for analyses.

7.3.2 Jet Calibration and Corrections

The goal of the jet calibration is to bring the energy and direction of the calorimeter jets found in the detector closer to the behavior of the MC truth particles without detector effects. Full calibration can be characterized by a procedure of about five steps. An overview of the complete chain can be seen in Fig 7.1. The steps are summarized as follows:

1. Origin Correction: A change in the jet direction to point back to the hard-scatter collision where the jet should originate from. By default, jet directions point to the center of the detector.

2. Pileup Correction: Pileup is all the extra energy signatures coming from other sources other than the hard-scatter event (defined in more detail later). Correcting for the contributions of pileup within a jet is done by subtracting terms related to the jet's area and the average number of pileup interactions

expected from the jet's energy.

3. MC Jet Energy Scale: A correction that is derived by computing the ratio of the energy of the truth particle jet to the reconstructed jet as a function of p_T and η.

4. Global Sequential Calibration: Using the track information combined with the geometrical location of the caloclusters, this correction reduces fluctuations and brings the data/MC closer together.

5. In-Situ Calibrations: This last correction is only applied to jets from actual data. It accounts for the differences between MC and data derived from taking the ratio of their energy.

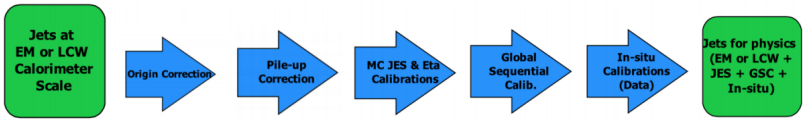

Figure 7.1: Overview of the jet calibration process, starting with raw energy clusters from the calorimeters to fully calibrated jets used in analyses.

Another correction to jets crucial especially for low p_T, is pileup removal.

7.3.3 Pileup (Removal)

As mentioned previously, pileup is the result of a high interaction environment that leads to signatures outside the hard-scatter event of interest. Since the LHC collides *bunches* of protons at high rates, the contributions from both in-time and out-of-time pileup are large. Both of these sources of pileup can contribute to additional calorimeter signals that can result in extra energy in jets of interest (handled by the calibration) as well as extra, low energy jets in the event considered. Dealing with these additional pileup jets is particularly challenging in low energy searches since the signal of interest can be washed away if the pileup removal is not handled properly.

There are a few variables designed to remove contributions from pileup: $corrJVF$ and R_{pT} [78]. Both are motivated by searching for tracks associated to the jet that originate from the vertex associated to the hard-scatter event of interest. $CorrJVF$ is defined as:

$$corrJVF = \frac{\Sigma_k p_T^{trk_k}(PV_0)}{\Sigma_k p_T^{trk_k}(PV_0) + \frac{\Sigma_{n>0}\Sigma_l p_T^{trk_l}(PV_n)}{(k \cdot n_{trk}^{PU})}} \tag{7.3}$$

where PV_0 refers to the hard-scatter vertex and PV_n refers to pileup vertices. From the equation, it can be seen that contributions from tracks originating from pileup decrease the jet corrJVF value (with a

maximum value of one). Thus, a lower $corrJVF$ value indicates a jet is more likely to originate from pileup. The other variable, R_{pT}:

$$R_{pT} = \frac{\Sigma_k p_T^{trk_k}(PV_0)}{p_T^{jet}} \tag{7.4}$$

is a measure of the energy contribution of the tracks from the hard-scatter vertex within the jet. Jets originating from pileup are more likely to have lower values of R_{pT}.

In order to maximize discrimination against pileup jets, a 2D likelihood called Jet Vertex Tagger (JVT) was constructed from both of the variables using controlled sample of jets. The JVT value of any jet can then be extracted from a jet's $corrJVF$ and R_{pT}, and be used to indicate if a jet originates from pileup or hard-scatter [78].

7.3.4 Jet Selection

Candidate jets are reconstructed from the calorimeter using the anti-k_t algorithm [77] with radius parameter $R = 0.4$. Jets are calibrated with the procedure described in the previous sections. No corrections for semi-leptonic b-hadron decays (resulting in muons that do not leave significant energy signatures in the calorimeter) are applied.

Jets are accepted with a kinematic selection of $p_T > 30\,\text{GeV}$ and $|\eta| < 2.5$. Quality criteria (also called jet cleaning procedure) are imposed to identify jets arising from non-collision sources or detector noise (using the LooseBad operating points) and any event containing at least one such jet is removed [79]. To reduce the contribution from jets associated with pileup, jets with $p_T < 60\,\text{GeV}$ and $|\eta| < 2.4$ are required to satisfy JVT > 0.59, where JVT is the output of the jet vertex tagger algorithm designed to discriminate between jets from the hard-scatter process and those from pileup [80].

To prevent double-counting of electron energy deposits as jets, jets within $\Delta R < 0.2$ of a reconstructed electron are removed. Finally, if the nearest jet surviving the above cut is within $\Delta R < 0.4$ of an electron, the electron is discarded, to ensure it is cleanly separated from nearby jet activity.

An important aspect for the H $\rightarrow$ aa $\rightarrow$ bbbb search is identifying the a-bosons. Since special jets are required for the search and considerable investigation was done selecting the proper candidate jets during the development of the double b-tagger, the exploration and discussion of these special jets are discussed in the following chapter about double b-tagging.

7.3.5 B-Tagging

The other important reconstruction technique for the analysis is finding jets that originate from b-quarks, known as b-tagging [81]. B-jets have special properties originating from the b-quark's unique physics that can be exploited to be experimentally distinguishable from other jets. B-quarks can decay into up or charm quarks via the weak force, but the decay lifetime is suppressed. This allows the b-hadrons to travel about 500 μm before decaying and creates a displaced *secondary vertex* inside the jet. As a result, there will be displaced tracks associated to the jet corresponding the the decay products of the b-quark. The decay chain of the b-hadron also leads to c-hadrons, which have a similar decay pattern (but less suppressed and a smaller lifetime) that can sometimes leave tertiary vertices. There are three algorithms that exploit these characteristics and serve as the baseline for b-tagging:

1. IP2D/IP3D: This algorithm uses the displaced tracks' impact parameter values to construct a log-likelihood. The impact parameters are defined with respect to the hard-scatter vertex, so larger values imply the tracks came from b-hadron decay products.

2. SV1: This method attempts to reconstruct the secondary vertex using the displaced tracks. This can be done by creating vertex candidates with all the track pairs in a jet. The vertex candidates are quality tested by reconstructing the mass to see if the vertex corresponds to uninteresting mesons and by checking the positioning of the vertex to ensure its origins are not from detector effects. The remaining two-track vertices are then combined into a single vertex. The mass and energy ratios of the tracks to this vertex are then used to build a discriminate.

3. JetFitter: This algorithm attempts to reconstruct the full decay chain of the b-hadron within the jet. Multiple vertices are looked for along the axis of the decay chain similarly to SV1 except the tracks are not combined into a single vertex. The mass and energy ratio are also computed, but now for the tracks corresponding to the full decay chain.

All of these algorithms capture important information individually, but the most widely used b-tagging algorithm in ATLAS combines variables from these algorithms with a multivariate technique to maximize discrimination of b-jets from jets originating from charm or the other lighter quarks (c-jets/light jets respectively) called MV2.

The MV2 algorithm is a boosted decision tree (BDT) trained to classify jets by flavor. Various MV2cxx classifiers exist in ATLAS with xx referring to the fraction of c-jets used during the training step of the BDT, e.g. MV2c10 is a BDT classifier trained with 90% light jets and 10% c-jets.

The MV2 tagger is calibrated at different working points (WP), which are defined by the acceptance of b-jets after cutting on a particular MV2 value. For example, at the 77% WP, 77% of b-jets are found with 95% purity, accepting 1 out of every 6 c-jets and 134 light jets. Multiple working points are defined to cover high efficiency/low purity performance and vice versa.

Single b-tagging is crucial to understanding the double b-tagger. Since the physics search relies on the dedicated double b-tagger developed for it, the next chapter will discuss all the aspects of the development and performance of the tagger.

Chapter 8

Low Mass Double B-Tagger

This chapter will describe the work done developing, optimizing, and calibrating a double b-tagger dedicated for low mass signatures. The data/MC samples will be described. The specific objects unique to the study will then be defined. After describing the training and optimization of the tagger, the calibration strategy will be detailed. The resulting corrections to MC bringing the behavior closer to real data are shown along with a description of the uncertainties in the measurements.

There are two major hurdles when defining low p_T double b-tagging. The first is defining a proper calibrated jet collection to reconstruct the a-bosons and Higgs mass properly for the intended physics search. The second is finding a method to identify the two b-quarks within the jet for proper tagging. The following section continues defining objects, focusing on the ones that are unique and/or developed for the purposes of the double b-tagger and addressing these problems.

8.1 Special Objects, Associations, and Flavor Labeling

8.1.1 Reclustered Calorimeter Jets

As described in section 1, the b-quarks of interest tend to be low p_T and collimated. The average ΔR between the b-quarks tends to be lower than $R = 0.4$. Thus, a larger collection of jets must be used. A previous study looking the MC jets with truth flavor showed that the optimal radius for signal efficiency was at $R = 1.0$ and $p_T > 30$ GeV. The current ATLAS standard $R = 1.0$ jets with trimming are only calibrated down to $p_T > 250$ GeV [82]. Thus, an alternative jet collection is required for the identification of a-bosons.

In order to have a jet collection that is calibrated to low enough p_T, calibrated anti-k_T $R = 0.4$ jets with the previously mentioned selection are reclustered with $R = 0.8$ [83, 84]. Since reclustered jets are reconstructed from calibrated anti-k_T $R = 0.4$ jets as their constituents, the reclustered jets will be properly calibrated and any jet uncertainties will propagate in a straightforward way. Thus providing a strong candidate jet collection for accurate mass reconstruction.

Reclustered jets have a selection of $p_T > 30$ GeV and pseudorapidity $|\eta| < 2.0$ applied.

8.1.2 Ghost Association

The other challenge to address is the method to double b-tag the reclustered jet collection. As described in section 7.3.5, tracks associated to jets are what are used to identify b-quark decay chains. By setting the distance parameter to $R = 0.8$ when reconstructing the reclustered jets, a larger radius for track association can be used. This allows more tracks from the targeted double b-quark decays to be associated to a reclustered jet.

The procedure for associating tracks to a given reclustered jets is done through ghost-association [85, 86]. By setting the p_T of the track 4-vectors to infinitesimally small values, the "ghost" tracks can then be reclustered with the constituents of the reclustered jets with the appropriate distance parameter. Since the p_T of the track 4-vectors is infinitesimally small, they will not influence the reconstruction of the jet. This method for track association is more robust in cases where ΔR track matching is ambiguous.

8.1.3 Exclusive-k_T Track Jets

Once a collection of tracks is associated to the reclustered jets, a method for identifying both b-quark decays must be employed. One way of doing this is by clustering the tracks in a way to yield track subjets. Ideally, the tracks from the same b-quark decays will be correlated with each other and the requirement of two subjets for a given reclustered jet should be the result from the two largest energy contributions in the jet (i.e. each b-quark).

The track jets are derived using the ghost-associated tracks to each reclustered jet as inputs to the exclusive-k_T (ExKt) method [87]. The tracks for a given jet are clustered together using the k_T algorithm. The k_T algorithm is very similar to the anti-k_T algorithm defined in section 7.3.1, except with k_T instead of $1/k_T$:

$$d_{ij} = \min\left(k_{ti}^2, k_{tj}^2\right) \frac{\Delta R_{ij}^2}{R^2} \tag{8.1}$$

$$d_{iB} = k_{ti}^2, \tag{8.2}$$

From this equation, it is evident that the k_T algorithm prefers to cluster soft objects first. The algorithm is set with a distance parameter of $R = 0.8$ to stop when there are exactly two track clusters left. Thus, the two remaining clusters will be approximately associated to the two highest energy contributions. These clusters are good candidates as the track jets associated to a given jet, which are referred to as "ExKt" track jets. Since the two remaining clusters should typically originate from the two highest energy contributions,

each ExKt track jet should have most of its energy coming from one b-quark for the jets with two b-quarks. In some cases, one of the subjets may contain the particles coming from the hadronization of both b-quarks. In this case, the exclusive-k_T algorithm is stopped when there are three subjets and the collection is used to try to identify the two b-jets and the radiation jet. Both subjet collections are used in the tagger and referred to as "ExKt2" and "ExKt3" subjets. All subjets used for the tagger studies have a selection of $p_T > 7$ GeV and pseudorapidity $|\eta| < 2.0$ applied.

8.1.4 Jet Flavor Labeling

The flavor label of MC reclustered jets is necessary to perform studies for the tagger efficiency. Flavor labels for reclustered jets are derived from flavor labels of the corresponding ExKt2 subjets. Flavor labeling of subjets is determined by first requiring $\Delta R < 0.8$ between a given weakly decaying b-hadron with $p_T > 5$ GeV and the parent reclustered jet of the subjet. For the pair of ExKt2 subjets in the parent jet matched a b-hadron, whichever subjet has the smallest ΔR with the b-hadron is labeled a b-jet. The process is repeated for c-hadrons and if a c-hadron is matched to a subjet that is not labeled as a b-jet, the subjet is labeled as a c-jet. If a subjet is not labeled as a b-jet or c-jet, it is labeled a light jet. The combined flavor labels for both subjets for a given reclustered jet determine the reclustered jet flavor label, which is described in more detail in section 8.5.

8.2 Data and MC Samples

For the development, optimization, and performance evaluation of the tagger, MC simulation samples were used. Since the task of developing a tagger essentially amounts to training a classifier (described in detail in section 8.3), an appropriate sample of "signal" and "background" jets is used. Jet samples were taken from MC samples used in the H $\rightarrow$ aa $\rightarrow$ bbbb analysis. Specifically, the $m_a = 20$ GeV signal (details described in section 6.2.1 and $t\bar{t}$ sample are used

A sample of $a \rightarrow$ bb (signal) jets was taken using signal samples of associated Higgs production with a W or Z boson. The samples with $m_a = 20$ GeV were used to ensure the b-quarks would be collimated enough to not be resolved into individual $R = 0.4$ jets. From the samples, jets labeled with 2 b-hadrons were used.

The "background" jets were taken from simulated semi-leptonic $t\bar{t}$ events. The $t\bar{t}$ sample was generated using POWHEG-BOX v2 [88] with the NNPDF3.0NLO PDF set. The ME to PS model parameter, h_{damp}, was set to $1.5m_t$. The PS and hadronization were modeled with Pythia8.210 with the A14 tune [89]. Jets with 1 b-hadron were used.

The truth labeled jets can be used as a labeled training set of data to that can be utilized for supervised learning methods. This class of learning methods employ statistical learning techniques to learn a mapping of input variables to a given label. Thus, the flavor labeled set of jets can be used to train a classifier that maps input variables to the flavor of a jet. This will be described more in the following section.

The calibration of the tagger uses gluons decaying into b-quark pairs (g→bb), which mimic the signal of interest (a → bb) and the 2015 pp collision data recorded by the ATLAS detector running at $\sqrt{s} = 13$ TeV with a total integrated luminosity of 3.2fb^{-1} [90]. The data used satisfies data quality requirements that ensure all events are collected under stable beam conditions with all detector subsystems operational. In addition to quality requirements, events must have a primary vertex that is defined with at least 3 tracks associated and with the largest $\sum p_T^2$ of all vertices. Events are selected with single jet triggers at various p_T thresholds. The HLT level jets are reconstructed with the anti-k_T algorithm with $R = 0.4$. The following jet trigger p_T thresholds were used: 15, 25, 35, 45, 55, 60, 85, 100, 110, 150, 175, 200, 260, 300, 320, 360 and 380 GeV. Many of these triggers are prescaled to reduce the large rate of data collection. The effect of this is discussed in section 8.5.

The MC samples used for the calibration of the tagger were the multijet MC samples. The samples were generated using PYTHIA 8 [91] with the LO NNPDF2.3 PDF set [92] with the A14 underlying event tune [89]. Multijet samples are produced at different various leading anti-k_T $R = 0.6$ truth jet p_T slices. In order to enrich the MC sample of g → bb jets, analogous truth muon-filtered multijet samples were used as well. The requirement of the truth muon enriches semileptonic b-decays, yielding more MC sample statistics with jets containing b-quarks. These samples were also produced with PYTHIA 8 with the same PDF set and tunings as the unfiltered samples.

8.3 Development of Low p_T Double B-Tagger

The low mass double b-tagger developed comes in the form of a BDT. It is similar to the idea of MV2c10 in the sense that a cut on the BDT output of the double tagger is used to classify bb-jets. In this case, the BDT is optimized to classify reclustered bb-jets against b-jets. As mentioned in section 8.2, the sample of reclustered jets was obtained from MC samples planned for use in the $H \to aa \to bbbb$ analysis. Specifically, jets with two b-hadrons are taken from the BSM $H \to aa \to bbbb$ sample (where $m_a = 20$ GeV) and jets with one b-hadron are taken from the $t\bar{t}$ sample. The b-hadron multiplicity labels the jet sample, with two b-hadrons in a jet being referred to as "signal" jets and one b-hadron in a jet as "background" jets.

8.4 Setup/Training

BDT training was done using the TMVA v4.2.0 and Root v5.36 software packages. Individual decision trees were trained by maximizing a decrease in the Gini index after each node split. A total of 500 trees each with maximum depth of three were trained using adaptive boosting (AdaBoost) with a learning rate of 0.5 [93].

From the original sample of signal and background reclustered jets, a training and testing set were created. Both training and testing samples have a mixture of jets with 1 or 2 b-hadrons with no overlap between samples. The actual training of the BDT was done using the training set, while the testing set was used to validate performance.

8.4.1 Feature selection

In order to find optimal features for the best classification performance, extensive studies were done to narrow down the set of input variables that would provide maximum discrimination. The driving force behind feature selection was simplicity and physics motivation. To reduce potential complications for the tagger calibration, track-based variables were emphasized as opposed to calorimeter-based variables. Specifically, variables were derived from track subjets obtained from running exclusive-k_T on the ghost-associated tracks of the reclustered jet. Since the b-hadrons in a given reclustered jet should contribute the majority of energy and tracks, the ExKt subjets reconstructed should correspond to the b-hadrons. From the set of variables related to track subjets, four were found to have the most discriminating power: the maximum and minimum MV2c10 values between the subjets, the ΔR between the subjets, and the p_T asymmetry of the subjets defined as $(p_T^1 - p_T^2)/(p_T^1 + p_T^2)$. As described in section 8.1, ExKt3 subjets are also created for a given reclustered jet to improve chances of having subjets with one b-hadron contained within them. Two sets of the aforementioned variables, one set from ExKt2 subjets and one set from the two highest MV2c10 valued ExKt3 subjets, are used for the BDT for a total of eight variables.

The minimum MV2c10 value between the subjets should generally be higher in reclustered jets with two b-hadrons since both subjets should have a b-hadron and subsequently have a higher score compared to a reclustered jet with one b-hadron. The maximum MV2c10 value simply adds b-tagging information about the other subjet, which may have correlations with the minimum valued subjet in reclustered jets with two b-hadrons that can be exploited from the BDT. Asymmetry in p_T should on average be less in reclustered jets with two b-hadrons since the b-hadrons should tend to split their energy/momentum evenly. The subjet ΔR in signal jets also tends to be lower, since the decay into the two b-hadrons tends to have them collimated, whereas the background jets have no such preference.

The distributions for the input variables comparing signal and background can be seen in Fig. 8.1. There are some additional structures in the background jet's subjet ΔR that can be attributed to the multiple sources of single b-jets.

8.4.2 Performance

A few common metrics were used to assess the performance of the tagger. To gauge how well the BDT classifies, a receiver operating curve (ROC) was produced and shown in Fig 8.2. This curve plots the background rejection as a function of the signal acceptance. Generally, as the background rejection increases to give higher signal purity, the signal acceptance decreases. Maximizing the background rejection and signal acceptance (or maximizing the area under the ROC) was a guiding factor in optimizing performance.

One general concern in statistical learning methods is overfitting a model to the statistical fluctuations of the training sample. This can come from a few different sources from training, but the effect is the same: reduced performance in samples outside of the training set. To assess overfitting, the BDT output distribution for both the training and test set were normalized and overlayed in Fig 8.2. Since the test set was not used in training the BDT, the performance is tested a completely independent sample from the training set with no influence on the BDT. If the training and test set BDT distributions are very similar, there is confidence that no overfitting has occurred.

The tagger efficiency was probed at two different working points (WP), corresponding to cutting on the BDT output above 0.3 (referred to as the tight WP) and between 0.3 and 0.1 (referred to as the loose WP). For the tight working point, the efficiency for jets with 2 b-hadrons is about 40%, accepting 1 in 20 jets with 1 b-hadron. For the loose working point, the efficiency for jets with 2 b-hadrons is about 25%, accepting 1 in 9 jets with 1 b-hadron.

Another assessment is in the correlations between variables. The correlations for signal and background jets can be seen in Fig 8.3. If correlations are extremely high between variables, it is an indicator that redundant information is being used in the BDT and increases the likelihood of overfitting. Most correlations between variables for both signal and background jets seem healthy. It should be noted that some higher correlations are expected due to the fact that there are "duplicate" sets of variables from two subjet collections (ExKt2 and ExKt3) that are not completely independent.

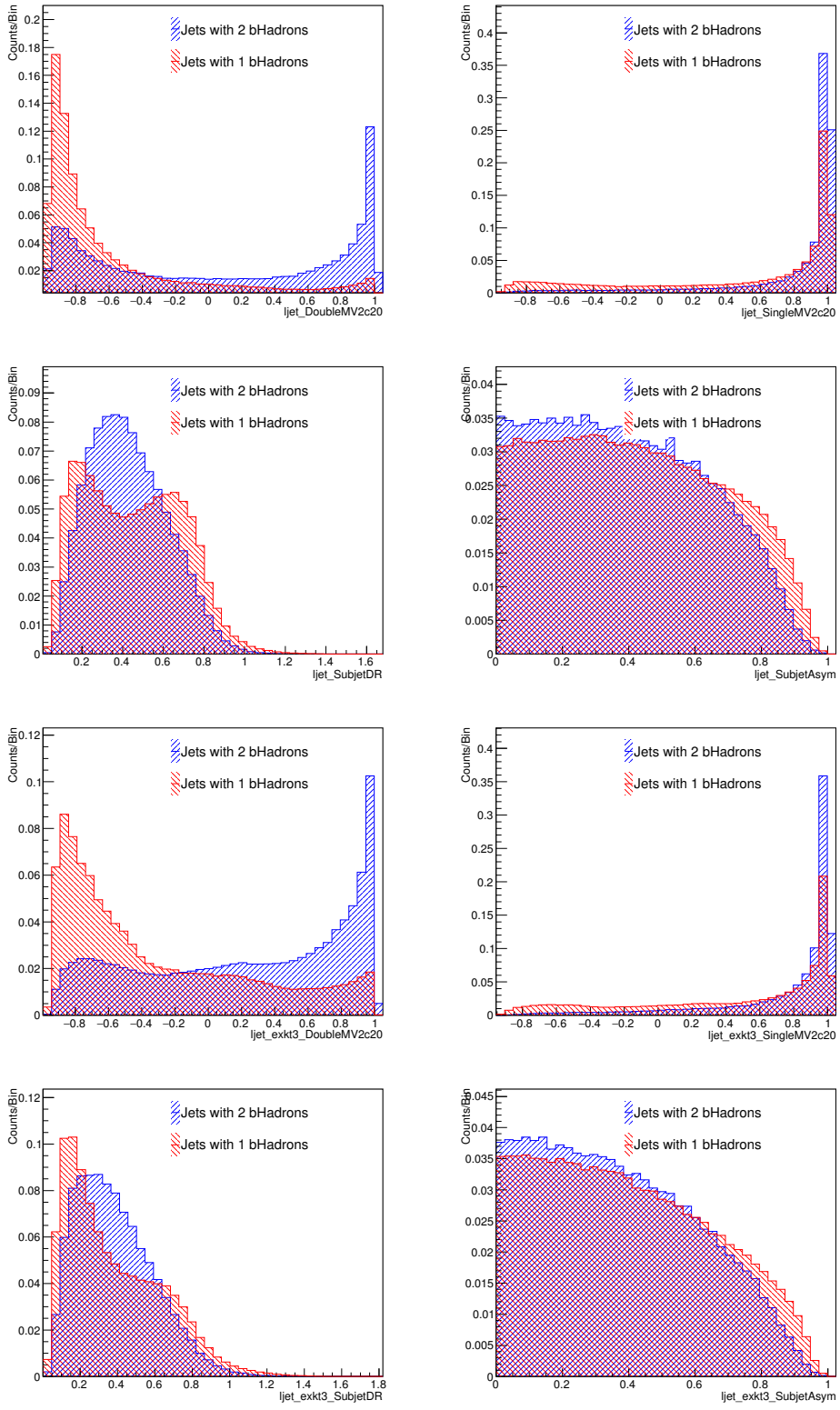

Figure 8.1: Inputs used in the preliminary BDT for the double b-tagger: Maximum MV2 value between subjets (left), minimum MV2 value between subjets (second from left), the subjet ΔR (second from right), and the subjet p_T asymmetry (right). The top row shows the distributions of from ExKt2 while the bottom shows the distributions from ExKt3.

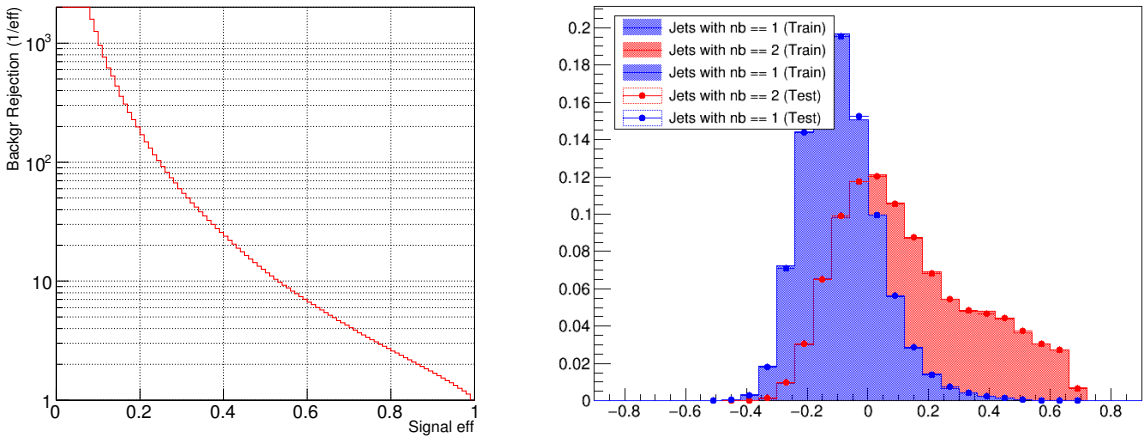

Figure 8.2: The Receiver Operating Curve (ROC) showing the signal acceptance against the background rejection (left). The BDT output for the training and test set for the signal and background jets (right).

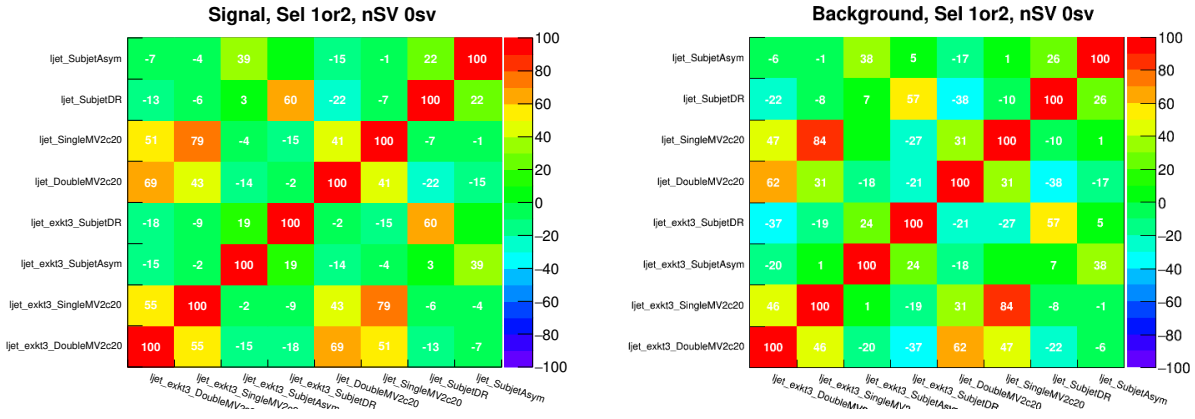

Figure 8.3: Correlations between the inputs to the BDT for the signal jets (left) and background jets (right). Identical variables from the two different subjet collections, ExKt2 and ExKt3, are expected to have higher correlations.

8.5 Tagger calibration using g→bb

The previous section dealt with understanding the performance of the tagger using MC simulation. Another crucial analysis is understanding the behavior in data and measuring any differences between them. This can be achieved by using a pure sample of bb-jets coming from Standard Model processes. The only reliable source for low p_T bb-jets comes from of g → bb splitting in QCD multi-jet events. Thus, the analysis for data/MC comparison will follow closely with what is done for the H → bb tagger validation [94] with modifications to suit the needs of the tagger.

8.6 Event selection

Data collected from single jet triggers was used. Since many of the triggers have prescales (in particular all jet triggers with $p_{T,\text{threshold}} < 360$ GeV), data samples were weighted event-by-event by a factor related to the prescale.

Selected events were required to have at least one $R = 0.8$ reclustered jet with $p_T > 30$ GeV and $|\eta| < 2.0$. The reclustered jets were considered as candidates for g $\rightarrow$ bb. To enrich the sample with bb-jets, exactly one of the subjets associated with gluon candidates was required to have a muon ΔR matched, which was labeled as j_{μ}^{trk} (the track subjet without the muon was labeled $j_{\text{non}-\mu}^{\text{trk}}$). To ensure the gluon candidate selection and the trigger selection were independent, the requirement that $\Delta R(j_{\text{trigger}}, j_{\text{gluon}}) > 1.5$ was imposed, where j_{trigger}, was defined as the leading, offline anti-k_T jet with radius $R = 0.4$.

8.7 Flavor fraction corrections

The flavor fractions are known to be mismodeled in the MC samples considered for g $\rightarrow$ bb comparison. In order to properly compare tagger efficiencies in data and MC, the mismodeling of the flavor composition must be corrected. To get the flavor fraction corrections, the MC was fit to the data using flavor sensitive distributions before tagging.

Due to the long decay length of heavy flavor hadrons, the impact parameter of the tracks associated to a jet lends itself as a good variable for flavor discrimination. In particular, the average signed impact parameter significance from the three highest p_T tracks for each track subjet for a given reclustered jet was used:

$$S_{d_0} = \frac{d_0}{\sigma(d_0)} s_j \tag{8.3}$$

where d_0 is the track transverse impact parameter with respect to the primary vertex, $\sigma(d_0)$ is the uncertainty on the d_0 measurement, and s_j is the sign of d_0 with respect to the jet axis, depending on whether the track crosses the jet axis in front of or behind the primary vertex. Some tracking detector effects are not taken into account in the MC simulation such as misalignment effects and missing simulation material. To accommodate the minor simulation mismodelings, the impact parameter distributions in MC were smeared as a function of p_T and η using Gaussian functions derived from 2015 minimum bias data.

The double flavor labeled S_{d_0} distributions for j_{μ}^{trk} and $j_{\text{non}-\mu}^{\text{trk}}$ were simultaneously used in a binned likelihood fit to derive the flavor fraction corrections. Since each of the two subjets can have flavor of b, c, or light, there are a total of 9 double flavor labels for a given reclustered jet. However, the flavor

combinations $(j_{\mu}^{\text{trk}}, j_{\text{non}-\mu}^{\text{trk}}) = \{(b,c),(c,b),(l,c),(l,b)\}$ were predicted to contribute very little to the total and were grouped into similar flavor categories. Specifically, the following flavor labels were merged together:

- $(l,l) = (l,l) + (l,c) + (l,b)$
- $(c,l) = (c,l) + (c,b)$
- $(b,l) = (b,l) + (b,c)$

Thus, the final flavor fraction components used in the fit were $(j_{\mu}^{\text{trk}}, j_{\text{non}-\mu}^{\text{trk}}) = \{(b,b),(c,c),(c,l),(b,l),(l,l)\}$. Since the flavor compositions and efficiency of the tagger can vary with the kinematics of the reclustered jet, the S_{d_0} templates are binned and analyzed as a function of reclustered jet p_T The bins used are j_{gluon} p_T : $\{(30-90),(90-140),(140-200)\}$ GeV.

The weights applied to the data from the prescale factors can sometimes be large ($> 10^5$). Since the study is binned in reclustered jet p_T and the trigger jet p_T is not one-to-one with the reclustered jet p_T, events with weights differing by large factors can enter the same bin. When building a likelihood (described in section 5) with events whose weights differ by several orders of magnitude, numerical instabilities occur that result in a fit that is not able to converge. In order to remedy the instability, an ad-hoc weight was applied to both data and MC depending on the leading, anti-k_T $R = 0.4$ jet p_T. Events with lower jet p_T were down-weighted more heavily than events with higher jet p_T. The reclustered jet p_T distributions are compared before and after applying the p_T dependent weight in Fig 8.4. After applying the weight, a discrepancy between data and MC simulation can be seen. The differences are understood to originate from the discrepancy in the offline small-R jet p_T spectrum between data and MC as shown in Fig 8.4.

Anti-k_T $R = 0.4$ Jet p_T, GeV	Weight
30-45	1/277503
45-55	1/75967
55-60	1/36985
60-85	1/25158
85-110	1/6168
110-150	1/2176
150-175	1/587
175-260	1/294
260-380	1/46
¿ 380	1

Table 8.1: Anti-k_T $R = 0.4$ jet p_T dependent weights applied to both data and MC to numerically stabilize the flavor fraction fit.

The inputs for the fit are then the weighted average S_{d_0} distributions of the track subjets. The binned

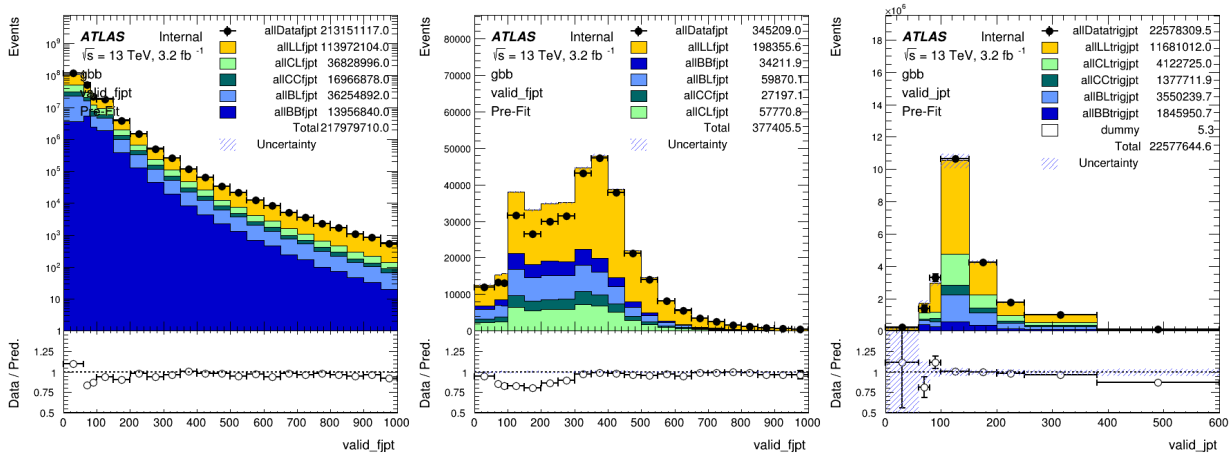

Figure 8.4: Large jet p_T spectrum after applying prescale factors to data (left). Large jet p_T spectrum after applying leading small-R jet p_T weights to both data and MC (center). Small-R jet p_T spectrum in the reclustered jet p_T range 50-200 GeV with prescale factors applied to data (right).

likelihood function is defined as:

$$\mathcal{L} = \prod_{i=1}^{n} p(y_i|\vec{\theta}) \tag{8.4}$$

where $\vec{\theta} \in \mathbb{R}^k$ is the parameter adjusting the normalization for each (k) flavor component, n is the number of bins used in the fit, and y_i is the number of data events in the i-th bin. The term, $p(y_i|\vec{\theta})$, is assumed to be a Poisson distribution:

$$p(y_i|\vec{\theta}) = \frac{(\vec{\theta} \cdot \vec{x}_i)^{y_i} \exp(-\vec{\theta} \cdot \vec{x}_i)}{y_i!} \tag{8.5}$$

with $\vec{x}_i$ being the vector of MC values for all flavor components in the i-th bin. The flavor fractions are determined by finding the $\vec{\theta}$ that maximizes $\mathcal{L}$. The binning used for the fit was chosen to minimize the number of bins to reduce the complexity of the fit while maintaining as much discriminatory power between the flavor components. The normalized flavor components used to determine the binning are shown in Fig 8.5. Since much of the discriminating power is in $S_{d_0} > 0$ and is held in lower S_{d_0} values, the binning was chosen to be (0, 2, 5, 10, 15, 20, 40, 80).

In order to assess the uncertainties on the flavor fraction corrections from the jet energy scale (JES) and track smearing, the MC samples were varied by 1 σ and re-fit for each systematic separately. The covariance matrix for each systematic is built assuming a correlation of 1 using the central values from the respective fits:

$$C_{ij} = (\theta_{\text{nom},i} - \theta_{\text{sys},i})(\theta_{\text{nom},j} - \theta_{\text{sys},j}) \tag{8.6}$$

where C_{ij} is the ij-th element of the covariance matrix indexed by flavor component and $\theta_{X,i}$ is the returned flavor component from the nominal or systematically varied fit. Each covariance matrix is summed together along with the statistical covariance matrix returned from the nominal fit. The diagonals of the total covariance matrix are taken as the uncertainties on each flavor component.

Post-fit S_{d_0} distributions for j_{μ}^{trk} and $j_{\text{non}-\mu}^{\text{trk}}$ are shown in Fig 8.6 and Fig 8.7. The post-fit flavor fractions are shown in tables 8.2, 8.3, 8.4.

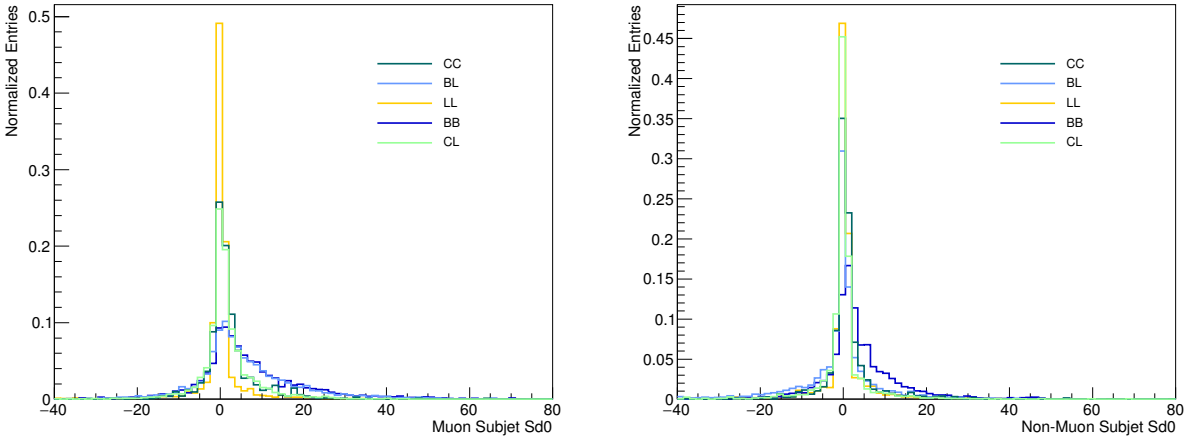

Figure 8.5: Normalized flavor Sd0 distributions for the muon (left) and non-muon (right) subjets.

	30 GeV $< p_T <$ 90 GeV $0.0 < \eta < 2.0$
BL	0.176 ± 0.004 (stat) ± 0.030 (sys)
BB	0.061 ± 0.004 (stat) ± 0.005 (sys)
LL	0.381 ± 0.011 (stat) ± 0.064 (sys)

Table 8.2: Flavor fractions post-fit for reclustered jet p_T bin 30 - 90 GeV.

	90 GeV $< p_T <$ 140 GeV $0.0 < \eta < 2.0$
BL	0.227 ± 0.004 (stat) ± 0.027 (sys)
BB	0.095 ± 0.001 (stat) ± 0.008 (sys)
LL	0.433 ± 0.016 (stat) ± 0.045 (sys)

Table 8.3: Flavor fractions post-fit for reclustered jet p_T bin 90 - 140 GeV.

	140 GeV $< p_T <$ 200 GeV $0.0 < \eta < 2.0$
BL	0.204 ± 0.003 (stat) ± 0.011 (sys)
BB	0.089 ± 0.001 (stat) ± 0.001 (sys)
LL	0.389 ± 0.012 (stat) ± 0.015 (sys)

Table 8.4: Flavor fractions post-fit for reclustered jet p_T bin 140 - 200 GeV.

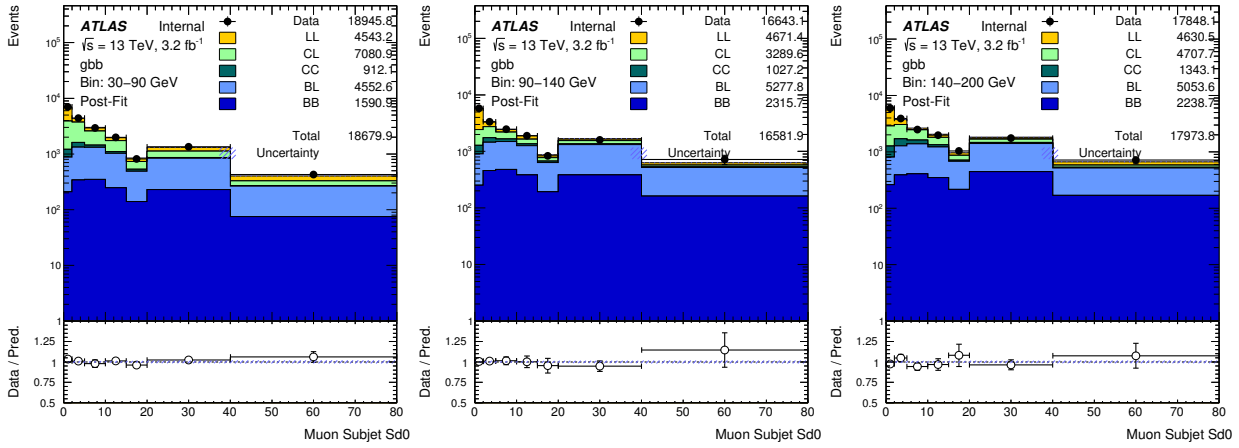

Figure 8.6: Post-fit muon subjet S_{d_0} distributions used for fit the flavor fraction compositions.

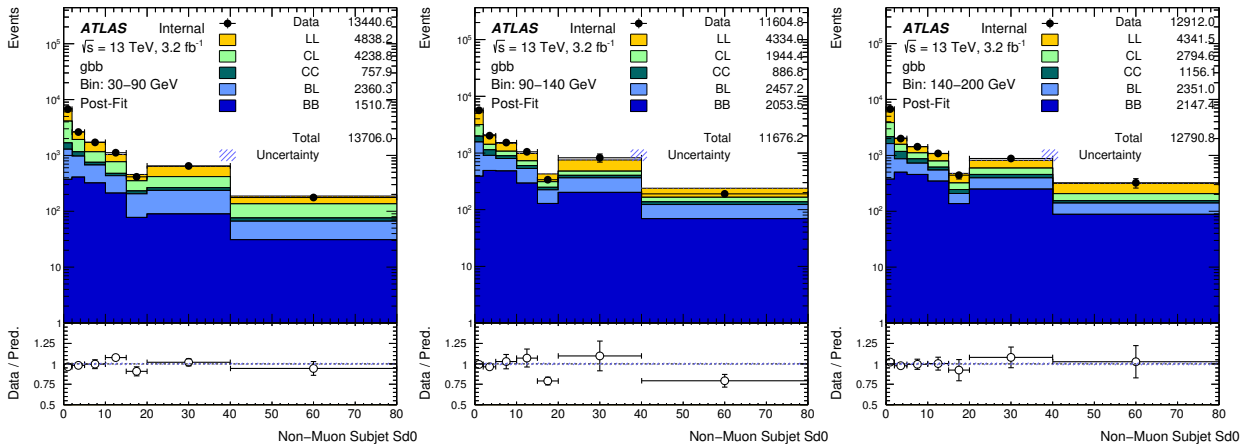

Figure 8.7: Post-fit non-muon subjet S_{d_0} distributions used for fit the flavor fraction compositions.

8.8 Double-Tagging Calibration Results

After applying the corrections to MC simulation, a data/MC comparison of the double tagger performance can be properly evaluated. The efficiency of the tagger depends on the flavor content of the reclustered jet. In particular, the $(j_\mu^{\text{trk}}, j_{\text{non}-\mu}^{\text{trk}}) = \{(b,b),(b,\ell)\}$ efficiencies are compared between data and MC simulation

in the form of a scale factor (SF):

$$SF_f = \frac{\epsilon_f^{\text{data}}}{\epsilon_f^{\text{MC}}} \tag{8.7}$$

where f denotes the flavor and ϵ is the efficiency for a given flavor in data or MC defined as the number of jets passing the BDT over the total number of jets.

In order to take into account the correlation between measuring the (b,b) and (b,l) flavor components, both were fit simultaneously in a binned likelihood fit, where the likelihood is once again assumed to be a product of Poisson distributions. To construct a sample that can directly fit the SF of interest and appropriately take the backgrounds $\{((\ell,\ell),(c,\ell),(c,c))\}$ into account, a bit of algebra is required to properly scale the MC and yield (b,b) and (b,ℓ). The statistical treatment for fitting of the scale factors (SF_{bb}, $\text{SF}_{b\ell}$) requires more care to construct a model to fit the SF directly. The data that satisfies the double b-tagging requirement can be separated into a background component, a bb component, and a $b\ell$ component.

$$N^{\text{data,post-tag}} = N_{\text{bkg}}^{\text{data,post-tag}} + N_{bb}^{\text{data,post-tag}} + N_{bl}^{\text{data,post-tag}} \tag{8.8}$$

In order to define the $N_{bb}^{\text{data,post-tag}}$ and $N_{b\ell}^{\text{data,post-tag}}$ components, the data is taken before double b-tagging and has the corrected double b-tagging efficiency applied:

$$N_{bb}^{\text{data,post-tag}} = (N_{bb}^{\text{data}}) \cdot (\epsilon_{bb}^{\text{data}}) = (N^{\text{data}} f_{bb}) \cdot (\epsilon_{bb}^{\text{MC}} \text{SF}_{bb}) = \left(N^{\text{data}} \frac{N_{bb}^{\text{MC}}}{N^{\text{MC}}} \right) \cdot \left(\frac{N_{bb}^{\text{MC,post-tag}}}{N_{bb}^{\text{MC}}} \text{SF}_{bb} \right) \tag{8.9}$$

where f_{bb} is the fraction of (b,b) jets before double b-tagging as measured in the flavor-fraction fit. The efficiencies are defined as:

$$\epsilon_{bb}^{\text{MC}} = \frac{N_{bb}^{\text{MC,post-tag}}}{N_{bb}^{\text{MC}}}; \qquad \epsilon_{bb}^{\text{data}} = \epsilon_{bb}^{\text{MC}} \text{SF}_{bb} \tag{8.10}$$

Simplifying equation (8.9), gives:

$$N_{bb}^{\text{data,post-tag}} = N^{\text{data}} \frac{N_{bb}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{bb} \tag{8.11}$$

A similar derivation for the $b\ell$ component yields:

$$N_{b\ell}^{\text{data,post-tag}} = N^{\text{data}} \frac{N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{b\ell} \tag{8.12}$$

The remaining piece is the background component $N_{\text{bkg}}^{\text{data,post-tag}}$. This component will not be corrected by

any scale factor and is defined using the corrected flavor fraction from simulation after double b-tagging:

$$\begin{aligned} N_{\text{bkg}}^{\text{data,post-tag}} &= N_{\ell\ell}^{\text{data,post-tag}} + N_{c\ell}^{\text{data,post-tag}} + \cdots \\ &= N^{\text{data,post-tag}} f_{\ell\ell}^{\text{post-tag}} + N^{\text{data,post-tag}} f_{c\ell}^{\text{post-tag}} + \cdots \\ &= N^{\text{data,post-tag}} \frac{N_{\ell\ell}^{\text{MC,post-tag}}}{N^{\text{MC,post-tag}}} + N^{\text{data,post-tag}} \frac{N_{c\ell}^{\text{MC,post-tag}}}{N^{\text{MC,post-tag}}} + \cdots \\ &= N^{\text{data,post-tag}} \left[\frac{N^{\text{MC,post-tag}} - N_{bb}^{\text{MC,post-tag}} - N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC,post-tag}}} \right] \end{aligned} \tag{8.13}$$

Adding equations (8.11), (8.12), and (8.13) together, a more operational version of equation (8.8) can be obtained:

$$\begin{aligned} N^{\text{data,post-tag}} = N^{\text{data,post-tag}} &\left[\frac{N^{\text{MC,post-tag}} - N_{bb}^{\text{MC,post-tag}} - N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC,post-tag}}} \right] \\ &+ N^{\text{data}} \frac{N_{bb}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{bb} \\ &+ N^{\text{data}} \frac{N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{b\ell} \end{aligned} \tag{8.14}$$

In order to build the likelihood function for SF_{bb} and $\text{SF}_{b\ell}$, $N^{\text{data,post-tag}}$ is solved using in equation (8.14):

$$\begin{aligned} N^{\text{data,post-tag}} \left[\frac{N_{bb}^{\text{MC,post-tag}} + N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC,post-tag}}} \right] &= N^{\text{data}} \frac{N_{bb}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{bb} + N^{\text{data}} \frac{N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{b\ell} \\ \Rightarrow N^{\text{data,post-tag}} = \left[\frac{N^{\text{MC,post-tag}}}{N_{bb}^{\text{MC,post-tag}} + N_{b\ell}^{\text{MC,post-tag}}} \right] &\cdot \left[N^{\text{data}} \frac{N_{bb}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{bb} + N^{\text{data}} \frac{N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC}}} \text{SF}_{b\ell} \right] \end{aligned} \tag{8.15}$$

The equation (8.15) allows one to define two background-subtracted components that are directly comparable to data:

$$\begin{aligned} N_{\text{MC,bkg subtracted}}^{\text{post-tag},bb} &= N^{\text{data}} \left(\frac{N^{\text{MC,post-tag}}}{N_{bb}^{\text{MC,post-tag}} + N_{b\ell}^{\text{MC,post-tag}}} \right) \left(\frac{N_{bb}^{\text{MC,post-tag}}}{N^{\text{MC}}} \right) \\ N_{\text{MC,bkg subtracted}}^{\text{post-tag},b\ell} &= N^{\text{data}} \left(\frac{N^{\text{MC,post-tag}}}{N_{bb}^{\text{MC,post-tag}} + N_{b\ell}^{\text{MC,post-tag}}} \right) \left(\frac{N_{b\ell}^{\text{MC,post-tag}}}{N^{\text{MC}}} \right) \end{aligned} \tag{8.16}$$

These background-subtracted versions encode not only the effect of $N_{bb}^{\text{MC,post-tag}}$ and $N_{b\ell}^{\text{MC,post-tag}}$ in the

signal, but also in the definition of the background fraction. The likelihood function used to fit is then a binned Poisson likelihood where each term is written as:

$$\mathcal{L}^I(\mathrm{SF}_{bb}, \mathrm{SF}_{b\ell}) = \mathcal{L}^I_{\mathrm{Poisson}}(N_I^{\text{data,post-tag}} | N_{bb,I}^{\text{MC,post-tag,bkg subtracted}} \mathrm{SF}_{bb} + N_{b\ell,I}^{\text{MC,post-tag,bkg subtracted}} \mathrm{SF}_{b\ell}) \tag{8.17}$$

where I runs over two S_{d_0} bins ($0 < S_{d_0} < 10\,\mu$m, $10 < S_{d_0} < 40\,\mu$m) in both the jet with a muon matched (j_μ^{trk}) and without it ($j_{\text{non}-\mu}^{\text{trk}}$). The resulting likelihood function is:

$$\mathcal{L}(\mathrm{SF}_{bb}, \mathrm{SF}_{b\ell}) = \mathcal{L}^{\mu}_{0<S_{d_0}<10} \cdot \mathcal{L}^{\mu}_{10<S_{d_0}<40} \cdot \mathcal{L}^{\text{non-}\mu}_{0<S_{d_0}<10} \cdot \mathcal{L}^{\text{non-}\mu}_{10<S_{d_0}<40} \tag{8.18}$$

The four bins are needed to provide discrimination between the (b,b) and (b,ℓ) components. The (b,b) component will be more prevalent for large S_{d_0} values. The fitted values of SF_{bb} and $\mathrm{SF}_{b\ell}$ are found by minimizing the negative logarithm of equation (8.18), similar to what was done in section 8.7. The resulting tight and loose SF with the uncertainties (described in the next section) are reported in Fig 8.8. After deriving the SF, they can be applied in equation (8.14) to compare the data post-tag to all the MC flavor components with corrected (b,b) and (b,ℓ) efficiencies. The post-fit, post-tag S_{d_0} distributions after applying SFs are shown in Fig 8.9 and 8.10 for tight and Fig 8.11 and 8.12 for loose.

8.9 Scale factor systematics

The following effects were considered for the uncertainty in the SF measurement: propagation of uncertainty from the flavor fraction measurement, impact of the jet energy scale uncertainties (JES), uncertainty in the track S_{d_0} smearing, and the impact of the bias from tagging a sample that has a muon.

The uncertainty from the flavor fraction measurement was assessed by using the total covariance matrix used for the flavor fraction measurement. Propagation of the uncertainty from the measurement was done by diagonalizing the covariance matrix and applying variations to the MC yields used for the SF measurement (pre-tag and post-tag):

$$N_{f,\text{varied}} = N_f \pm \sum_{m=1}^{5} \lambda_m (\text{Eigenvector}_m[f]) N_f \tag{8.19}$$

The varied (b,b) and (b,ℓ) post-tag samples were fit to get the impact on the SF measurement. Additionally, 1 σ variations from JES and track systematics were applied directly to the post-tag (b,b) and (b,ℓ) samples separately and fit to assess their direct impacts on the SF measurement.

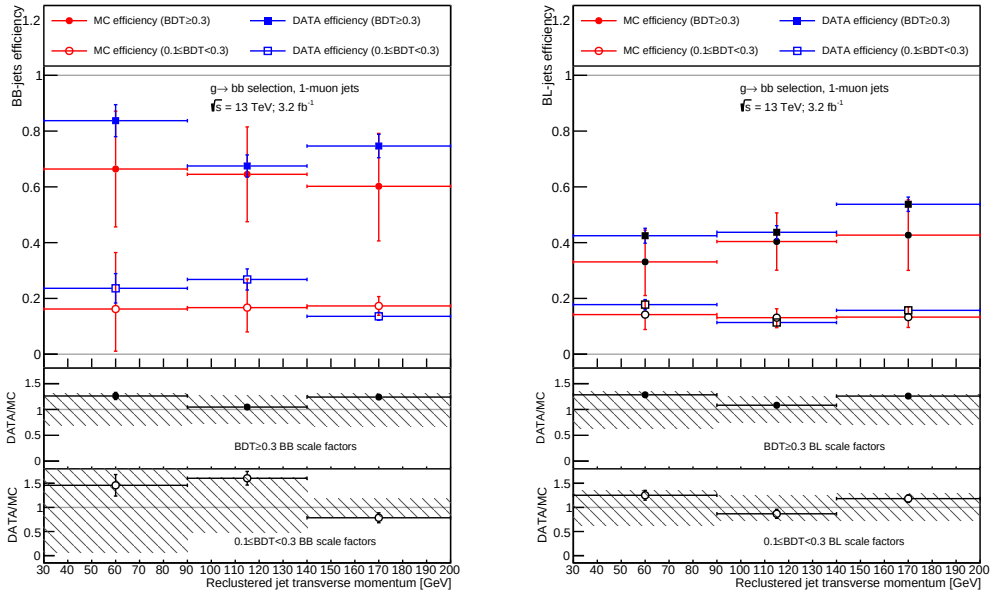

Figure 8.8: Data and MC efficiencies reported for each reclustered jet p_T bin. The (b,b) component is on the left and the (b,l) component is on the right. Complete uncertainties are reported. The scale factors can directly be read off of the ratio plot.

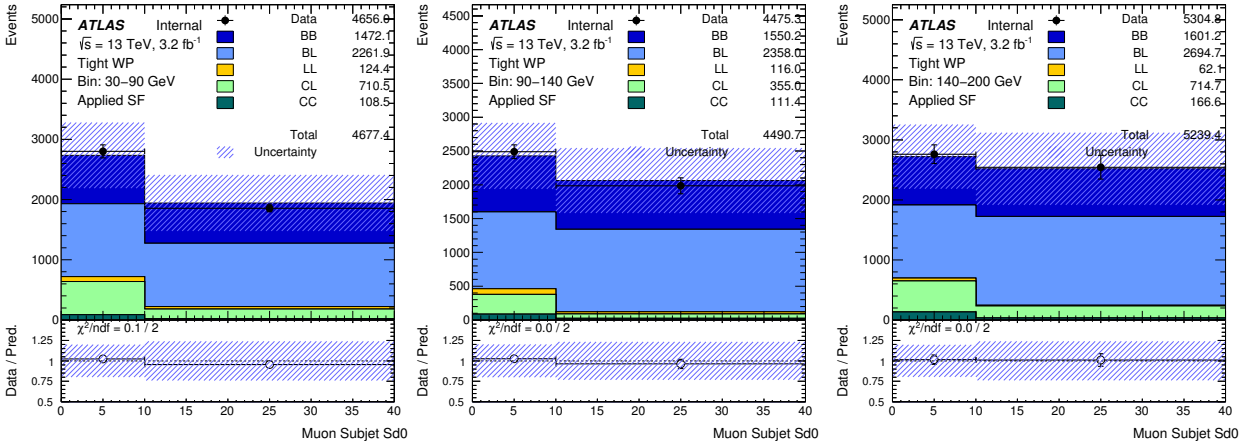

Figure 8.9: Post-fit S_{d_0} distributions from subjets with a muon used to fit the Scale Factors in the sample of reclustered jets with a muon in exactly one subjet for the tight WP.

The uncertainty from the bias from requiring a muon was estimated by repeating the SF measurements using a selection that required both subjets to have a muon ΔR matched. Since both subjets have a muon, the physics motivation for splitting the S_{d_0} distributions by muon and non-muon subjet does not apply. Both subjet's average S_{d_0} are combined into a single S_{d_0} distribution.

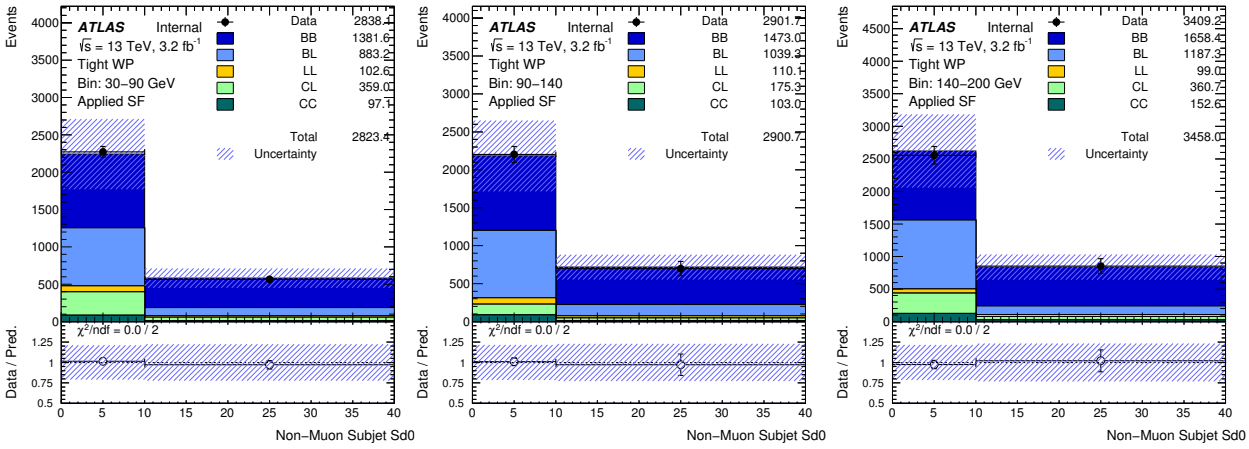

Figure 8.10: Post-fit S_{d_0} distributions from subjets without a muon used to fit the Scale Factors in the sample of reclustered jets with a muon in exactly one subjet for the tight WP.

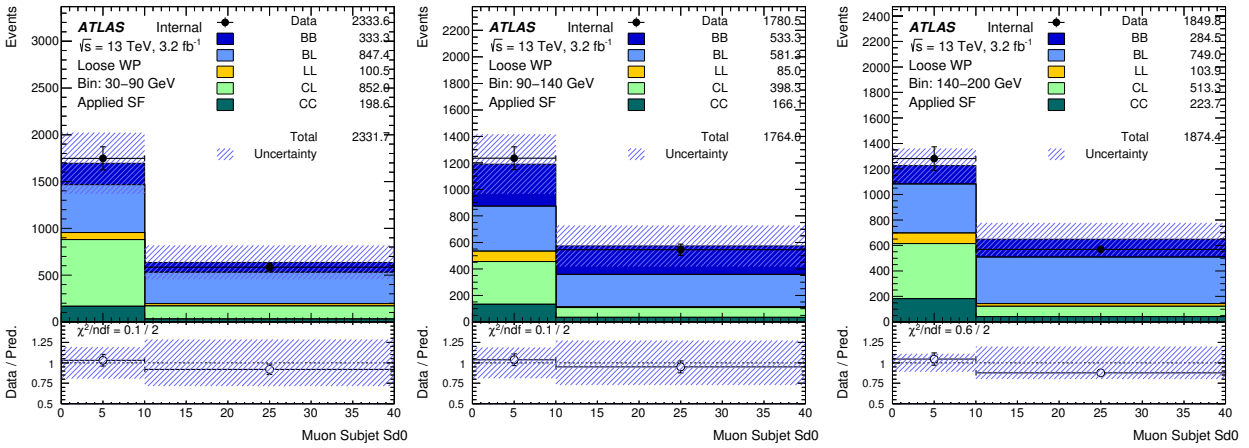

Figure 8.11: Post-fit S_{d_0} distributions from subjets with a muon used to fit the Scale Factors in the sample of reclustered jets with a muon in exactly one subjet for the loose WP.

The measured flavor fractions described in section 8.7 are applied to the 2 muon sample. The post-tag distributions are constructed as described in section 8.8 using the tight WP, but the (b, ℓ) component is not considered because it is too small in a 2 muon, post-tag sample. The (b, b) SF measurement is performed with a likelihood fit using the flavor corrected, post-tag combined S_{d_0} sample. The average fractional difference between the tight WP (b, b) SF from the 1 (nominal) and 2 muon samples across all reclustered jet p_T and η bins was then taken as the uncertainty due to the bias and applied to both the (b, b) and (b, ℓ) for both WP as uncertainty. Post-tag, post-fit S_{d_0} distributions and subsequent SF are shown in Fig 8.13 and 8.14, respectively.

A binned likelihood is performed in the 2 muon sample with the same binning and likelihood assumptions

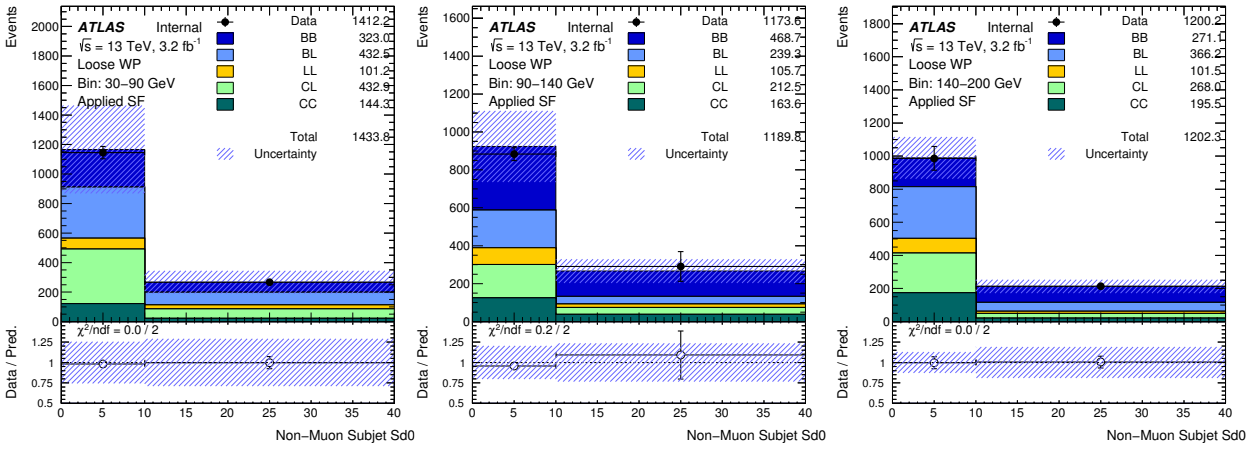

Figure 8.12: Post-fit S_{d_0} distributions from subjets without a muon used to fit the Scale Factors in the sample of reclustered jets with a muon in exactly one subjet for the loose WP.

as the nominal fit.

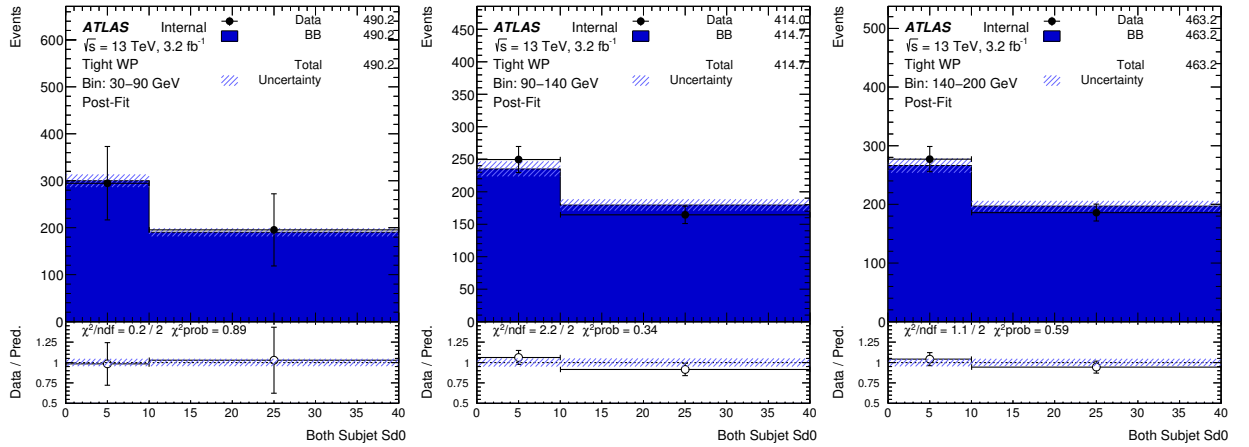

Figure 8.13: Post-fit S_{d_0} distributions from both subjets used to fit the Scale Factors in the sample of reclustered jets with a muon in each subjet.

8.10 Additional Studies

8.10.1 Light jet muon fakes

A concern in the measurement of the flavor fraction corrections is the potential mismodeling due to muon "fakes" in the light flavor jets. Since light jets have no heavy flavor decays that can semileptonically decay into muons, the light jets "fake" muons originating from pion decays. To cross-check if the fakes caused

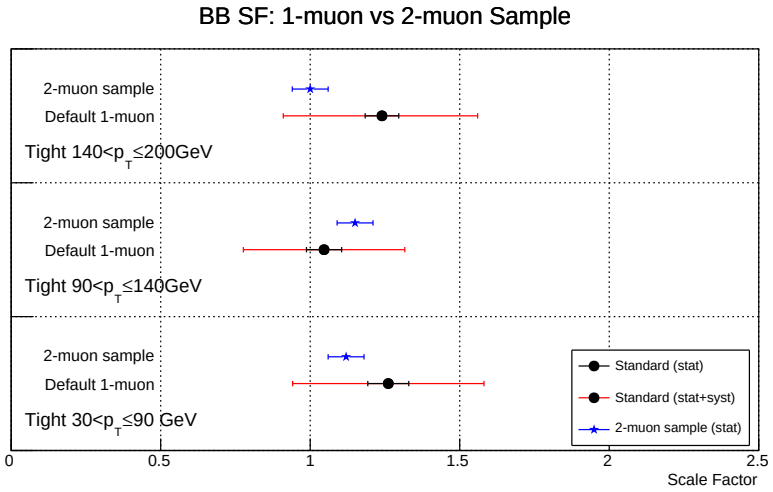

Figure 8.14: Scale factors reported for each reclustered jet p_T and η bin from the sample of jets with a muon in each subjet.

mismodeling in the flavor fraction corrections, the measurement was redone using a variable sensitive to identifying fakes. The sample of jets that had a muon in each subjet was used in order to have a pure sample of jets with decays-in-flight.

Muons are reconstructed with a combination of measurements from the inner detector (ID) and muon spectrometer (MS). Due to the nature of decays-in-flight, the independent measurements in the ID and MS will have differences. A variable sensitive to the differences is the p_T asymmetry of the two measurements:

$$\rho = \frac{p_T^{MS} - p_T^{ID}}{p_T^{ID}} \tag{8.20}$$

where p_T^{MS} is the transverse momentum measured in the MS corrected for proper extrapolation back to the IP and corrected for energy loss in the calorimeter.

Two separate flavor fraction fits were done and compared to assess the effect of the fakes. The first is a fit to the S_{d_0} distributions from the sample of jets with 2 muons. Since both subjets have muons, the distinction between muon and non-muon subjet is not necessary. The combination of both subjet S_{d_0} values is combined into one distribution fit for the flavor fractions.

The other flavor fraction fit uses two regions were defined by a cut on ρ. A sample of jets with 2 muons is used again, but with different distributions being fit. A "symmetric" region was defined as S_{d_0} values coming from jets with muons that satisfied $0.95 < \rho < 1.05$. Any jets with muons that satisfied ρ outside of the window were classified into the "asymmetric" region and should be enriched with fakes. Both regions were then simultaneously fit using a binned likelihood with corrections to the flavor components being fit

140 GeV $< p_T <$ 200 GeV $0.0 < \eta < 2.0$	
postfit BB (standard) postfit BB (ρ split)	0.36 ± 0.03 0.39 ± 0.03
postfit BL (standard) postfit BL (ρ split)	0.18 ± 0.02 0.23 ± 0.02
postfit LL (standard) postfit LL (ρ split)	0.14 ± 0.02 0.19 ± 0.02

Table 8.5: Flavor fractions post-fit for the ρ split fit and the nominal 2 muon sample fit.

by maximizing the likelihood. The comparison between the corrections found with ρ and the 2 muon fit be found in table 8.5. Both measurements yield the same corrections within 1σ, indicating that the fakes are not impacting the measurement.

8.10.2 Differences between a→bb and g→bb

Since the jet sources differ between the training (a→bb) and calibration (g→bb) samples, differences in the input distributions for the tagger may affect the scale factors which can be seen in figures 8.15 and 8.16. Kinematic differences may arise due to the fact that the a-boson is massive, unlike the gluon. While these differences can result in changes in the efficiency of the tagger for a given flavor, it is the data/MC comparisons of efficiency (SF) that are important to investigate between the processes. Ideally, the calibration would be done for a process similar to a→bb, such as Z→bb or H→bb. However, these processes do not samples of low p_T, sufficiently collimated b-quarks. In order to assess the effect from the different jet sources, SF were derived using a reweighted g→bb post-tag MC sample fit to the nominal g→bb post-tag MC sample.

The reweighting was derived using the double tagger input variables sensitive to the kinematic differences, the subjet ΔR and p_T asymmetry. Taking the normalized input distributions for jets from a→bb and g→bb, the ratio of the two was applied as a weight. This was done for both the subjet ΔR and p_T asymmetry, giving two weights that both get applied to each (b, b) and (b, ℓ) reclustered jet in the g→bb sample. With the reweighted sample, the fitting procedure described in 8.8 was done. Instead of fitting data to MC as before, the data was replaced with the reweighted sample. Any SF deviations from 1 are attributed to differences in the kinematics between the jet sources. The (b, b) and (b, ℓ) SF with statistical errors are reported in 8.17. All values are consistent with 1 within statistical error except the lowest p_T and η bin, but this is also consistent when compared with the uncertainties to the nominal measurement (i.e. a deviation of 0.14 in the (b, b) SF is well covered by the uncertainty).

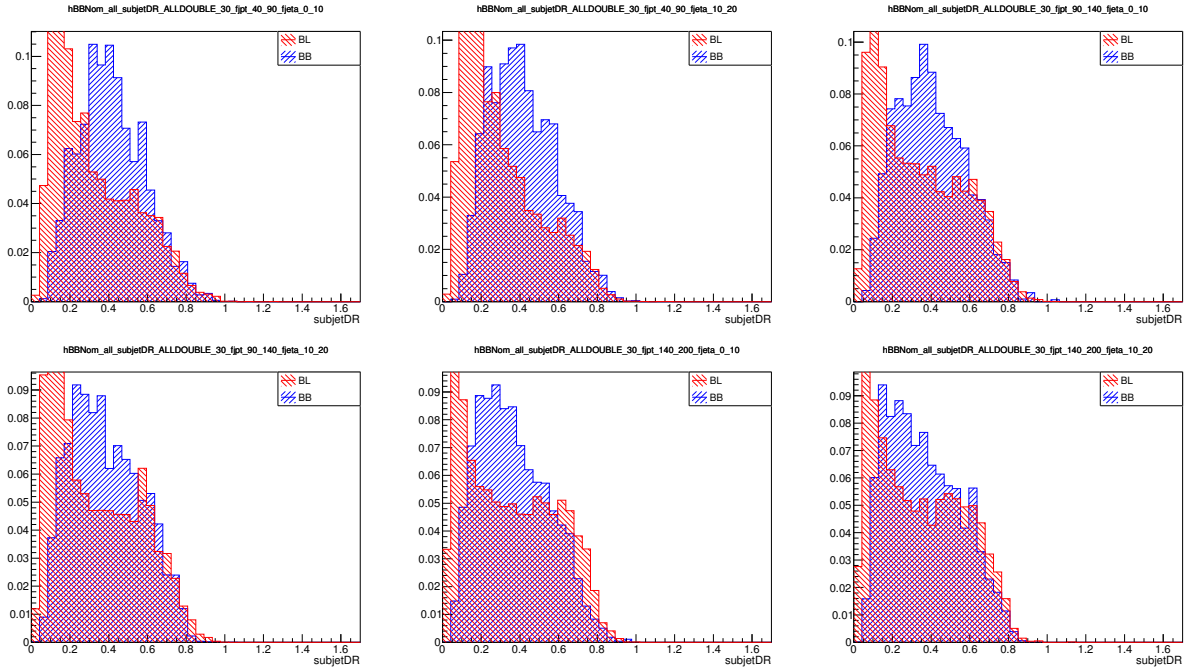

Figure 8.15: ExKt2 subjet ΔR from gluons for each reclustered jet p_T and η bin.

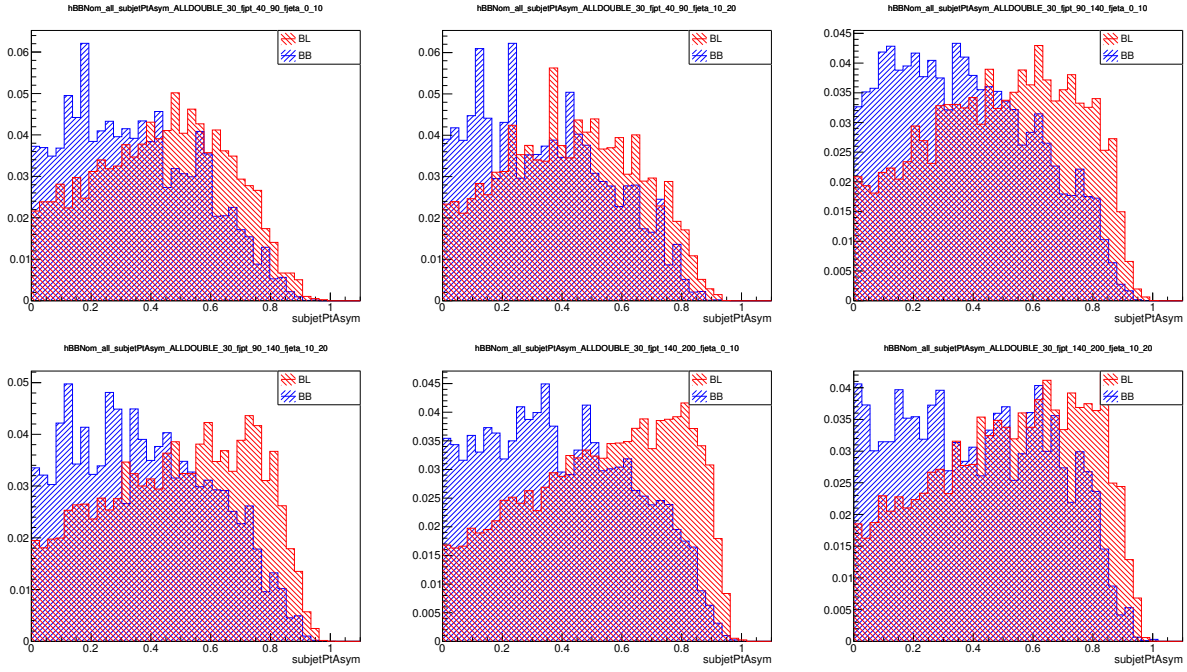

Figure 8.16: ExKt2 subjet p_T asymmetry from gluons for each reclustered jet p_T and η bin.

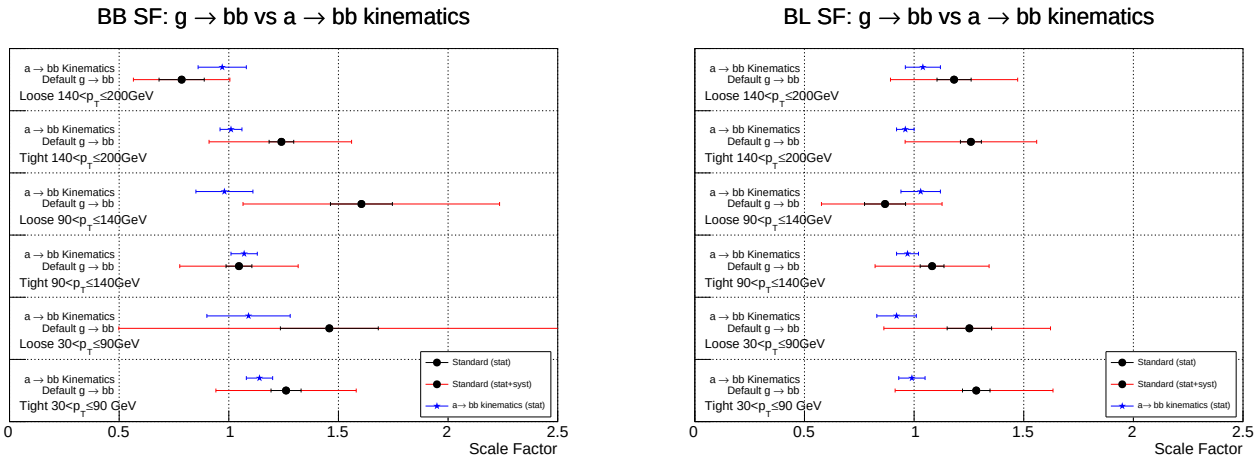

Figure 8.17: Scale factors from comparing a→bb to g→bb reported for each reclustered jet p_T bin for (b,b) (left figure) and (b,l) (right figure) flavors for both WP. Statistical uncertainties are reported.

8.11 Conclusion

Development and calibration for a new double b-tagger dedicated to low mass searches has been presented. The tagger discriminant is the result of training a BDT classifier on MC samples relevant to the searches that will employ the tagger, in this case exotic Higgs and $t\bar{t}$ samples. The classifier was optimized to distinguish between double and single b-jets. Double b-jet candidates were constructed by reclustering $R = 0.4$ anti-k_T jets with $R = 0.8$. Variables related to the kinematics and b-tagging information for the exclusive-k_T track subjets built from associated tracks were fed as inputs to the tagger.

In order to calibrate the tagger performance, data/MC comparisons were done using g→bb to derive scale factors that correct the MC efficiency. Since the dijet MC samples used in the calibration are known to have flavor mismodelings, a fit to the subjet S_{d_0} distributions was done to correct the flavor composition of the sample. After applying the flavor corrections, the post-tag (b,b) and (b,ℓ) flavor components were fit to data to yield the aforementioned scale factors. Uncertainties on the scale factors were considered from jet energy scale, track smearing, propagation of uncertainty from the flavor composition correction measurement, and bias due to requiring a muon in the g→bb candidates. The following chapter will describe the $H \to 2a \to 4b$ strategy with the application of the tagger.

Chapter 9

Search for H → aa → bbbb

After full development and calibration of the tagger, it was then applied to the $H \to 2a \to 4b$ low mass search. This chapter outlines the full analysis strategy and results obtained searching for the exotic Higgs decay.

9.1 Data sample and event preselection

This search is based on pp collision data at $\sqrt{s} = 13$ TeV and with 25 ns bunch spacing collected by the ATLAS experiment in 2015 and 2016. Only events recorded with a single-electron or single-muon trigger under stable beam conditions and for which all detector subsystems were operational (in particular the toroid and IBL are on) are considered. The corresponding integrated luminosity uses the 2015 data sets (periods D-J5), corresponding to 3,212.96 pb, and the 2016 data sets (periods A3-L), corresponding to 32,861.6 pb. The total integrated luminosity for the 2015+2016 data sets is therefore 36,074.56 pb.

Single lepton triggers with different p_{T} thresholds are combined in a logical OR in order to increase the overall efficiency. When using the 2015 data set, the following trigger chains are used:

- electrons: `HLT_e24_lhmedium_L1EM20VH OR HLT_e60_lhmedium OR HLT_e120_lhloose`,
- muons: `HLT_mu20_iloose_L1MU15 OR HLT_mu50`.

When using the 2016 data set, the following trigger chains are used:

- electrons: `HLT_e24_lhtight_nod0_ivarloose OR HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e140_lhloose_nod0`
- muons: `HLT_mu24_ivarmedium OR HLT_mu26_ivarmedium OR HLT_mu50`

The triggers with the lower p_{T} threshold include isolation requirements on the candidate lepton, resulting in inefficiencies at high p_{T} that are recovered by the triggers with higher p_{T} threshold.

For simulated data, a pseudo run-number is randomly assigned to each event. The pseudo run-numbers are taken among the 65 runs of the 2015 data set and the 150 runs of the 2016 data set, in a way to give the simulated events with a given run-number the same integrated luminosity as the corresponding real data run. Simulated events with a pseudo run-number corresponding to the 2015 (2016) data set are required to satisfy the same trigger requirements as the real data events of the 2015 (2016) data set.

Events satisfying the trigger selection are required to have at least one reconstructed vertex with at least five associated tracks with $p_{\mathrm{T}} > 400$ MeV, consistent with originating from the beam collision region in the x–y plane. If more than one vertex is found, the hard-scatter primary vertex is taken to be the one which has the largest sum of the squared transverse momenta of its associated tracks.

Events are required to have exactly two reconstructed electrons or muons. Since single lepton triggers are used, one of the two reconstructed leptons, $m_{\ell\ell}$, is required to match within $\Delta R < 0.15$ the lepton reconstructed by the trigger. The leptons are required to be oppositely signed and their invariant mass is required to be within 81 GeV $< m_{\ell\ell} <$ 101 GeV to identify the Z boson and suppress the $t\bar{t}$ background.

9.2 Signal and background modeling

After the event preselection, the main background is from $t\bar{t}$+jets and Z+jets production. Backgrounds containing at least two prompt leptons other than $t\bar{t}/Z$+jets production arise from diboson, Wt-channel single top-quark production, and $t\bar{t}V$ processes. There are also several processes which may contain non-prompt leptons that satisfy the lepton isolation requirements or jets misidentified as leptons. The background contribution from misidentified leptons is small after the requirement of at least two b-jets.

As in the case of the signal samples, Monte Carlo background samples also include the simulation of pileup interactions, and are processed through a full GEANT4 detector simulation and the same reconstruction software as the data. Further details about the modeling of each of the backgrounds are provided below.

9.2.1 Signal modeling

The signal sample described in section 6.2.1 and 6.2 is used for the signal modeling. The Higgs with m_H = 125 GeV is produced in associated with a Z boson and decays into two a-bosons, which each decay in a pair of b-quarks. As shown in section 6.2.1, below m_a = 30 GeV is the threshold where the b-quarks start becoming too collimated to resolve and require a different strategy to identify the $a \to bb$ resonances. Thus, the mass points of m_a = 15, 20, 25 GeV were used for the dedicated low mass search search.

9.2.2 $t\bar{t}$+jets background

The $t\bar{t}$ sample described in section 8.2 is used for the background modeling. For the double b-tagging studies, this sample was used to retrieve a sample of flavor labeled reclustered jets. In the analysis, additional treatments are applied to the sample to properly use it in the analysis. More details will be described as necessary.

The nominal $t\bar{t}$+jets sample is generated using the Powheg-Box v2 NLO generator with the NNPDF3.0 PDF set [95]. The *hdamp* parameter, which controls the p_{T} of the first additional emission beyond the Born configuration, is set to 1.5 times the top-quark mass of $m_t = 172.5$ GeV. The main effect of the hdamp setting is to regulate the high-p_{T} emission against which the $t\bar{t}$ system recoils. Parton shower and hadronization are modeled by Pythia 8.2 with the A14 [96] tune. The sample includes event weights so that we can evaluate uncertainties from the choice of scales, the choice of hdamp, and the parton distribution function used. The sample is generated separately for "non-all hadronic" W boson decays, dilepton decays, and all-hadronic decays. The dilepton sample is a subset of the first sample. Additional samples for each of these types of W boson decays are generated with filters requiring extra b-jets, which is important to ensure a reasonable number of simulated events passing cuts in the high b-tag multiplicity region.

The EvtGen v1.2.0 [97] program is used to simulate the bottom and charm hadron decays. The sample is normalized to the top++2.0 [98] theoretical cross-section of 832^{+46}_{-51} pb, calculated at next-to-next-to-leading order (NNLO) in QCD that includes resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [99–103].

Small improvements to the modeling are made by reweighting $t\bar{t}$ samples to NNLO predictions [104, 105], in particular for the top-quark p_{T}. A reweighting of the p_{T} of the $t\bar{t}$ system was also studied (as it was done in [106]), but the simulation samples were found to be in good agreement with the NNLO predictions such that this extra reweighting is not needed.

The $t\bar{t}$+jets samples are generated inclusively, but events are categorized depending on the flavor of partons that are matched to particle jets that do not originate from the decay of the $t\bar{t}$ system. Particle jets are reconstructed from all stable truth particles (not counting muons and neutrinos) with the anti-k_t algorithm with a radius parameter $R = 0.4$ and are required to have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$. Events where at least one such particle jet is matched within $\Delta R < 0.4$ to a b-hadron with $p_{\mathrm{T}} > 5$ GeV not originating from a top-quark decay are labeled as $t\bar{t}+b\bar{b}$ events. Similarly, events not previously categorized as $t\bar{t}$+$b\bar{b}$, and where at least one particle jet is matched within $\Delta R < 0.4$ to a c-hadron with $p_{\mathrm{T}} > 5$ GeV not originating from a W boson decay, are labeled as $t\bar{t}+c\bar{c}$ events. Events labeled as either $t\bar{t}+b\bar{b}$ or $t\bar{t}+c\bar{c}$ are generically referred to below as $t\bar{t}$+HF events, where HF stands for "heavy flavor". The remaining events

are labeled as $t\bar{t}$+light-jet events, including those with no additional jets. This classification is made for the purpose of comparisons to other $t\bar{t}$+jets event generators and the propagation of systematic uncertainties related to the modeling of $t\bar{t}$+HF.

The additional heavy flavor jets produced in association with the $t\bar{t}$ pair in POWHEG-BOX+Pythia8 are modeled by the backward evolution of the parton shower. Detailed studies of the theoretical uncertainties associated with this description were performed in Run 1 using the categorization of $t\bar{t}$+HF. In particular, the POWHEG-BOX+Pythia8 sample was compared with a dedicated ME+PS matched $t\bar{t}$+jets sample generated with MadGraph5 1.5.11 using the CT10 PDF set with tree-level diagrams of up to three additional partons (including b- and c-quarks) and interfaced to Pythia6 .425. The studies showed that the description of extra radiation in the Matrix Element description is in reasonable agreement with the parton shower model used. More details of this study can be found in App. G of Ref. [107].

The modeling of the $t\bar{t} + b\bar{b}$ background, particularly important for the search described in this note, is improved by reweighting the POWHEG-BOX+Pythia8 prediction to an NLO prediction of $t\bar{t} + b\bar{b}$ including parton showering [108], based on Sherpa+OpenLoops [109, 110] using the CT10 PDF set. This reweighting is performed for different topologies of $t\bar{t} + b\bar{b}$ in such a way that the inter-normalization of each of the categories and the relevant kinematic distributions are at NLO accuracy. In particular, the relative cross-section in each category is adjusted to the NLO prediction. In addition, reweightings are performed to take into account kinematic differences between Powheg and NLO $t\bar{t}+b\bar{b}$, based on the p_{T} of the top-quark and $t\bar{t}$ systems and on the p_{T} and η of the heavy flavor jet in the topologies with only one additional heavy flavor jet. The reweighting improves the modeling for a range of kinematic variables used to validate the reweighting. Note that the inclusive $t\bar{t} + b\bar{b}$ cross-section is not changed by the reweighting. More studies on $t\bar{t}$ modeling and categorization can be found in Ref. [111].

9.2.3 W/Z+jets background

Samples of W/Z+jets events are generated with the Sherpa v2.2.1 [109] generator. The ME calculation is merged with the Sherpa PS [112] using the ME+PS@NLO prescription [113]. The PDF set used for the ME calculation is CT10 with a dedicated PS tuning developed by the Sherpa authors. Both the W + jets and Z + jets samples are normalized to their respective inclusive NNLO theoretical cross-section calculated with FEWZ [114]. Samples are produced for each of the W + jets and Z + jets categories using filters for a b-jet (Wb or Zb+ jets), a c-jet and no b-jet (Wc or Zc+jets), and with a veto on b and c-jets (W or Z+light-jets).

9.2.4 Other simulated background

Samples of single top-quark backgrounds corresponding to the Wt and s-channel production mechanisms are generated with POWHEG-BOX v1 [115, 116] using the CT10 PDF set. Overlaps between the $t\bar{t}$ and Wt final states are removed using the "diagram removal" scheme [117]. Samples of t-channel single top-quark events are generated using the POWHEG-BOX (ST_tchan_4f) [118] NLO generator that uses the 4-flavor scheme (4FS). The single top quark samples are normalized to the approximate NNLO theoretical cross sections [119–121].

Samples of diboson production $WW/WZ/ZZ$+jets are generated with the NLO generator SHERPA v2.1.1 and include processes containing up to four electroweak vertices. The matrix element includes 0 partons at NLO and up to 3 partons at LO using the same procedure as for W/Z + jets. The final states simulated require one of the bosons to decay leptonically and the other boson to decay hadronically. All diboson samples are normalized to their NLO theoretical cross-sections provided by SHERPA.

Samples of $t\bar{t}V$ events, including $t\bar{t}WW$, are generated with up to two additional partons using MADGRAPH5 and interfaced to Pythia8 with the A14 NNPDF23LO UE tune. A sample of $t\bar{t}H$ events is generated with the MG5_aMC@NLO generator using the CT10 PDF set. Showering is performed using Herwig++ and the UE-EE-5 tune with the CTEQ6L1 PDF set. Inclusive decays of the Higgs boson are assumed in the generation of the $t\bar{t}H$ sample.

9.3 Analysis strategy

As described previously, this search focuses on the $H \to 2a \to 4b$ process, where the H boson is produced in association with a Z boson. At low m_a, the b-quarks from the same parent a are expected to be collimated. Therefore, the signal is expected to have two leptons from the Z decay as well as two double b-tagged jets coming from the two a-bosons decaying from the H boson. Events are categorized into signal and background regions.

9.3.1 Signal and background region definitions

After applying the preselection, events that passed having exactly two double b-tagged reclustered jets were investigated. Both tight and loose working points for double tagging were used to probe for signal and control regions. Several kinematic distributions were explored to optimize separation of signal and background.

9.3.1.1 Reclustered jet mass difference

The double b-tagged reclustered jets coming from the signal H will be reconstructed a-bosons. Thus, the mass of each double tagged jet should be equal to m_a. Since it is assumed that both a-bosons decaying from the same H should have the same mass, the difference in mass between both double tagged jets should be approximately 0. Double tagged reclustered jets originating from background processes will come from additional gluon radiation $g \to bb$, coincidental overlap of single b decays, or mistags of single b-jets. None of these conditions put a requirement of the mass of the jet or constrain the difference between a pair of double tagged jets. Figure 9.1 shows the distribution of the mass difference for all expected processes.

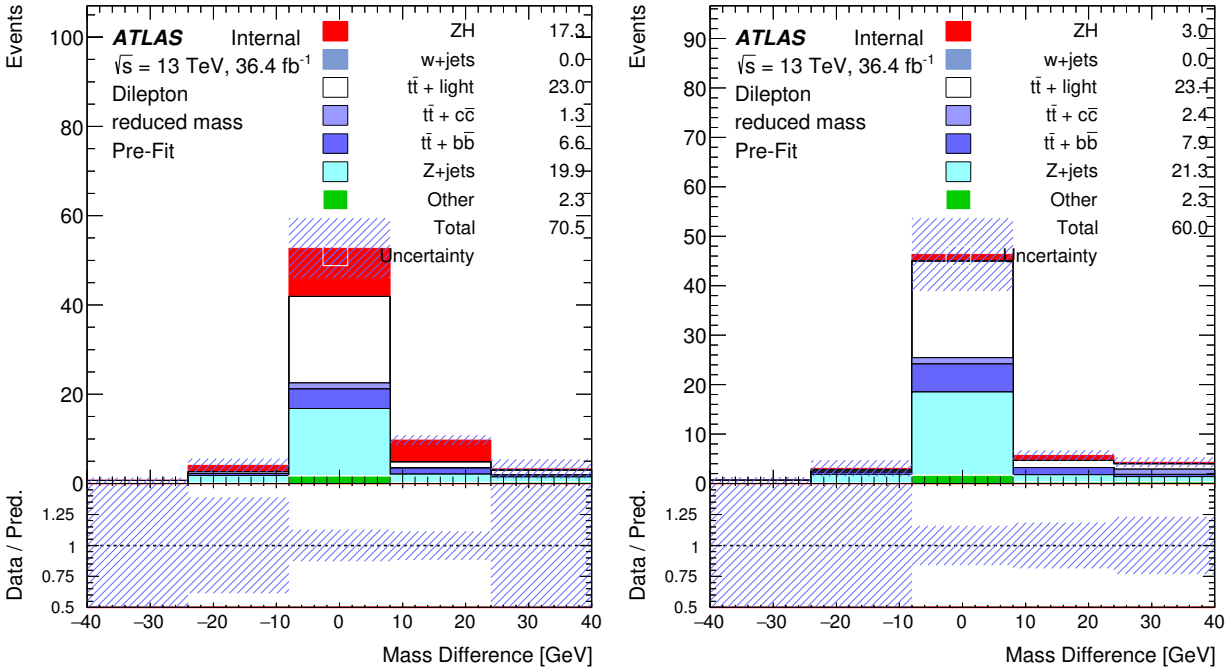

Figure 9.1: Distribution of mass difference between two tightly double b-tagged reclustered jets for signal and background processes after applying the preselection. The 20 GeV m_a hypothesis is shown on the left and the 30 GeV m_a hypothesis is shown on the right.

9.3.1.2 Reclustered jet pair mass reconstruction

For the signal, two a-bosons are originate from the Higgs boson. Thus, a pair of double b-tagged reclustered jets will have a combined mass around $m_H = 125$ GeV. While the b-quarks in the double tagged jets for background processes have no such constraint. Since the sources for double tagged jets vary for background processes, no structure is expected for a mass pair and it should be a good discriminatory variable.

The discrimination efficacy can be improved by reducing the width of the H mass distribution. The

individual reclustered jet masses will often fluctuate around m_a. By subtracting them from the pair mass, the fluctuations can be removed and width of the pair mass can be reduced. This variable is referred to as the "reduced mass". The improvement in mass width can be compared in figure 9.2. Figure 9.3 shows the distribution compared for all expected processes in the form:

$$m^{\text{reduced}}_{(bb)(bb)} = m_{(bb)(bb)} - 125 \text{ GeV} - (m^1_{bb} - m_a) - (m^2_{bb} - m_a) \tag{9.1}$$

where $(bb)(bb)$ denotes the pair of double tagged jets and m^x refers to the mass of the individual jets and m^a is the mass of the signal a under a given mass hypothesis.

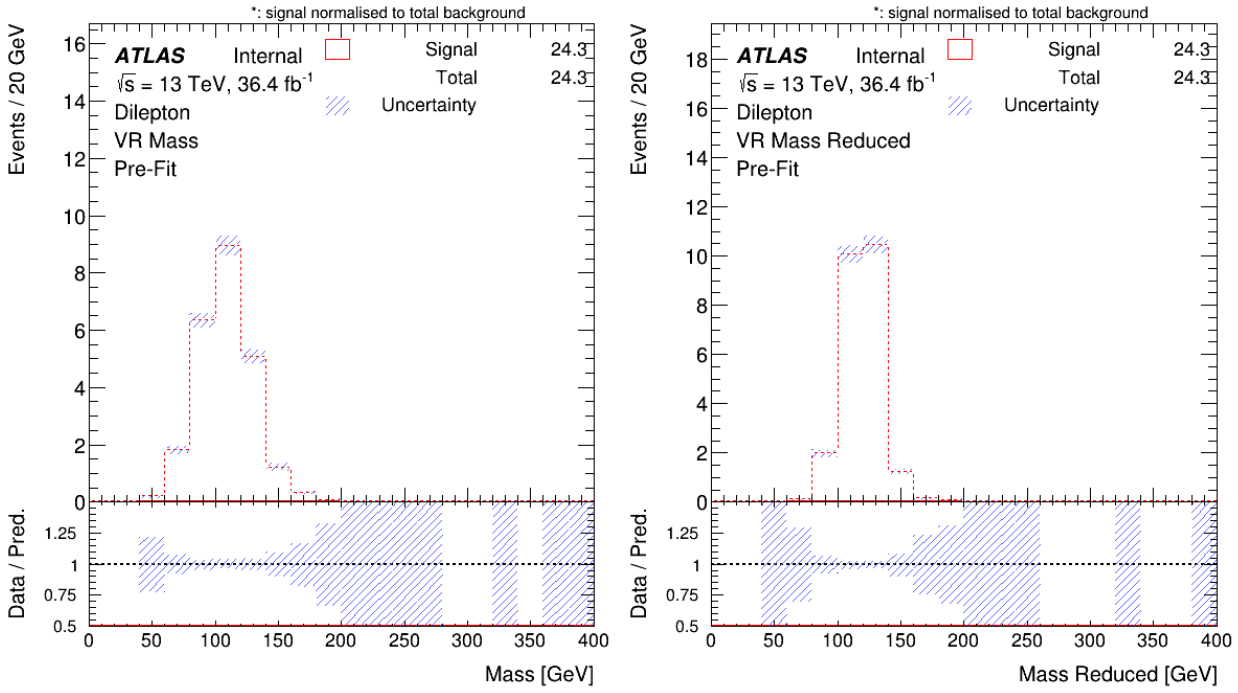

Figure 9.2: Mass (left) and reduced mass (right) distribution for the pair of two tightly double b-tagged reclustered jets for signal and background processes after applying the preselection for the 20 GeV mass hypothesis.

9.3.1.3 Application of double b-tagger SF

Scale factors are applied on a reclustered jet-by-jet basis for each event in such a way to reweight the event based on the likelihood of tagging each jet. Specifically, the probability for each event in MC is:

$$P(\text{MC}) = \prod_{i=\text{tight}} \epsilon_i \prod_{i=\text{loose}} \epsilon_{\text{loose},i} \prod_{i=\text{neither}} (1 - \epsilon_i - \epsilon_{\text{loose},i}) \tag{9.2}$$

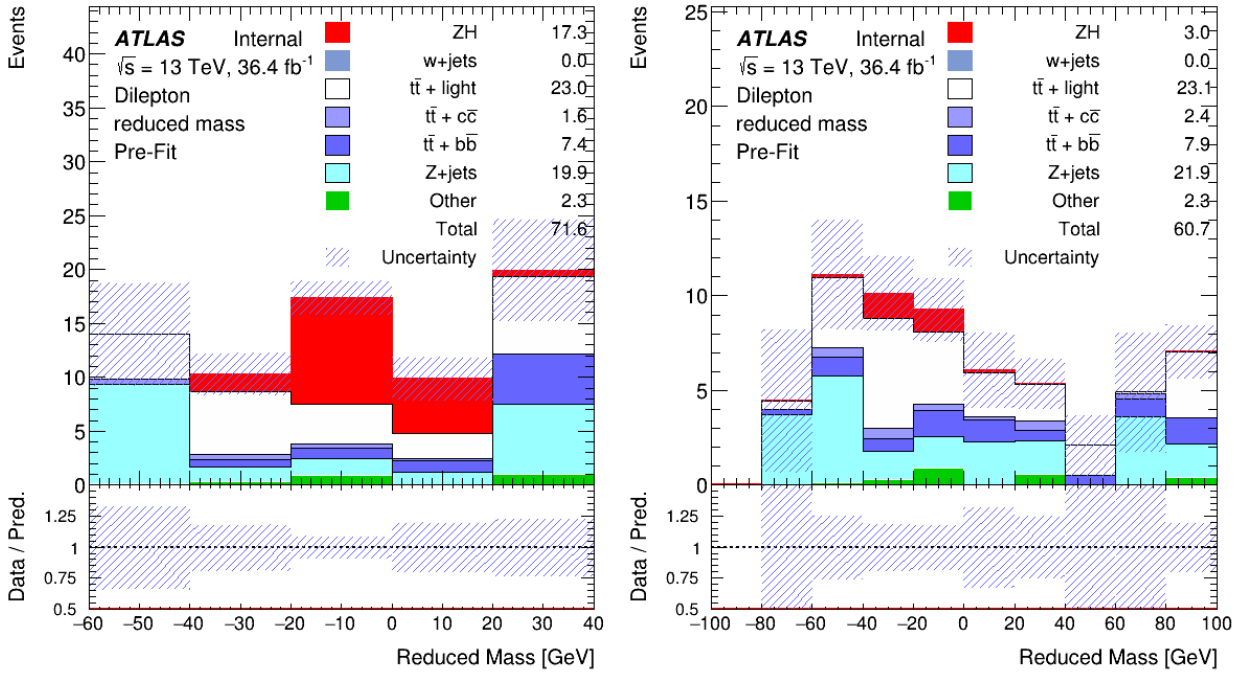

Figure 9.3: Distribution of mass for the pair of two tightly double b-tagged reclustered jets for signal and background processes after applying the preselection. The 20 GeV m_a hypothesis is shown on the left and the 30 GeV m_a hypothesis is shown on the right.

and for data:

$$P(\text{data}) = \prod_{i=\text{tight}} \text{SF}_i\epsilon_i \prod_{i=\text{loose}} \text{SF}_{\text{loose},i}\epsilon_{\text{loose},i} \prod_{i=\text{neither}} (1 - \text{SF}_i\epsilon_i - \text{SF}_{\text{loose},i}\epsilon_{\text{loose},i}) \tag{9.3}$$

with the weight applied to each MC event given by $w = P(\text{data})/P(\text{MC})$. All signal and control region figures in the following sections have the double tagging weight applied.

9.3.1.4 Signal regions

After the preselection, cuts based on the aforementioned variables were also applied. Figure 9.4 shows a 2-dimensional plot of the mass difference and reduced mass. The cuts were defined from a window that captures the most of the signal, while reducing the background as much as possible. The optimal cuts on the reduced mass will vary depending on the signal mass point being probed since the jet masses being subtracted out will change depending on the mass hypothesis. For now, the 20 GeV point is considered for the rest of this chapter. The 30 GeV mass point is not included due to the low signal yield seen in Fig 9.3. This is expected due to the fact that the higher mass results in less collimated b-quarks and higher jet

multiplicity. The previous resolved analysis targeted the higher masses.

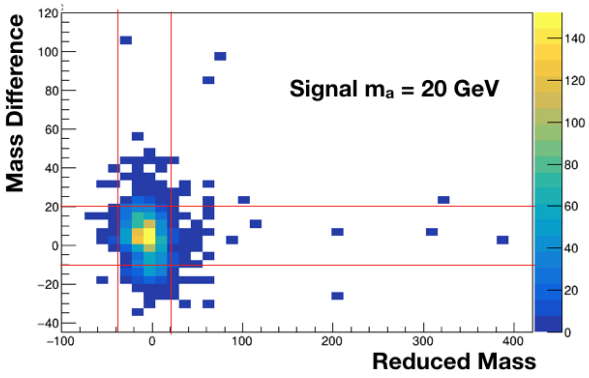

Figure 9.4: 2D plot of the reduced mass plotted against the mass difference for the 20 GeV mass point signal.

Generally, the cuts have the form:

$$20 \text{ GeV} + 2m_a > m_H^{\text{reduced}} - 125 \text{ GeV} > -40 \text{ GeV} + 2m_a \tag{9.4}$$

$$25 \text{ GeV} > m_{\text{diff}} \text{ GeV} > -25 \text{ GeV} \tag{9.5}$$

where m_a is taken to be 20 GeV since this did not affect the distribution significantly for the different mass points.

One signal region was defined using the preselection with the tight WP definition for double b-tagging and the mass window cut. Another signal region was defined with the same cuts as before, only with the requirement of one loose and one tight double b-tagged jet. The following section will describe the motivation for considering a loose WP.

Other regions were considered in order to calibrate the main background predictions and constrain the relevant systematic uncertainties.

9.3.1.5 Control regions

For the $t\bar{t}$ background, in most cases the events contain two true leptons from leptonic decays of the two W bosons, and the two b-jets from the top quark decays satisfy the b-tagging requirements. Additional jets arise from gluon splitting into $b\bar{b}$ and $c\bar{c}$ and from additional jets from initial state radiation and pile-up. Events will enter the signal regions from mistagging single b-jets as double b-jets, double tagging jets that have overlapping b's from different sources, and/or double tagged jets from gluon splitting. In both signal regions, events with mistagged single b-jets comprise the majority of the $t\bar{t}$ background contribution.

For the Z+jets background, the dominant contribution is from Z bosons produced in association with gluon splitting into $b\bar{b}$. The Z+jets background events enter the signal regions through either actual double b-tagged jets originating from the gluon splitting or mistagged single b-jets also originating from gluon splitting, but resolved into individual jets. The majority of Z+jets events entering both signal regions come from the mistags.

In order to define regions enriched with the major backgrounds, two requirements were adjusted: the double tag working point and dilepton mass cut. Relaxing the working point on the double b-tagger will enrich the mistag rate and result in regions with higher samples of mistagged single b-jets which are the dominant backgrounds in the signal regions. The two tight double b-tagging requirement was relaxed to have at least one loose tagged double b-jet, resulting in events with one tight and one loose double b-jet or two loose double b-jets. In order to have control regions enriched with $t\bar{t}$, all previous regions defined with the various additional jet and double b-tagging requirements were investigated looking outside of the defined Z mass window.

9.3.1.6 Validation regions

Additionally, validation regions were defined to assess the systematic uncertainties. These regions have no impact on calibrating the background modeling, but are instead used to compare MC predictions and data post-fit to cross-check the level of agreement. The first region is defined within the Z mass window, but outside of the reduced mass cut ($m_H^{\text{reduced}} - 125\ \text{GeV} > -40\ \text{GeV} + 2m_a$). The particular region was chosen to be enriched with Z+jets to ensure the background modeling is adequate. Another region was chosen that has all the same requirements as the two tight tagged signal region, except the leptons are required to be opposite flavored (one electron and one muon). This region was chosen to have a sample with the other dominant background, $t\bar{t}$.

9.3.1.7 Summary of regions

Figure 9.5 shows the MC simulated predicted yields for each process in the signal, control and validation regions. Figure 9.6 shows the relative composition of the backgrounds in all the regions, with $t\bar{t}$ and Z+jets as the main backgrounds.

A statistical analysis based on a likelihood function constructed as a product of Poisson probability terms over all regions in single bins considered in the search is used to derive the background predictions and uncertainties, and to test for the presence of a signal. The statistical procedure is described in Section 9.5, and the results are shown in Section 9.6.

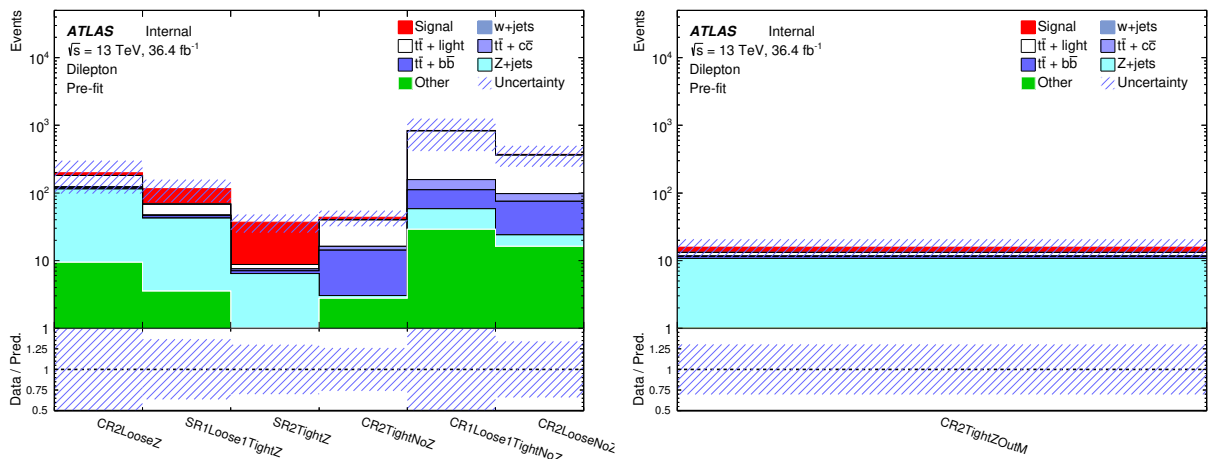

Figure 9.5: Signal and control regions (left) and validation regions (right) for the analysis showing the background composition and the expected signal for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$.

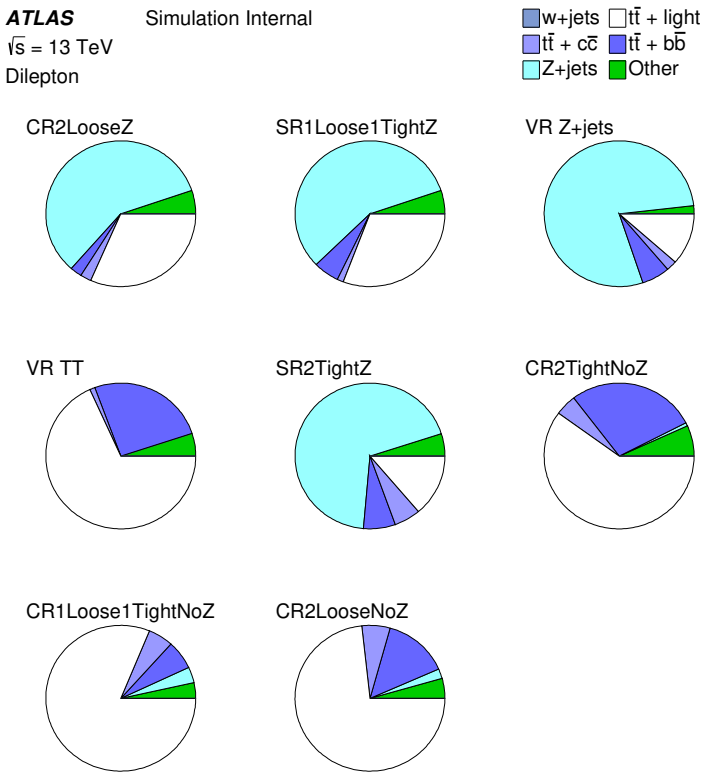

Figure 9.6: Pie charts showing the background compositions in all regions.

9.4 Systematic uncertainties

Several sources of systematic uncertainty are considered that can affect the normalization of simulated samples and/or the relative amounts in the signal and control regions. Individual sources of systematic uncertainty are considered uncorrelated. Correlations of a given systematic uncertainty are maintained across processes and channels. Table 9.1 presents a list of all systematic uncertainties considered in the analysis and indicates whether they affect only normalization or both shape and normalization. The full correlation matrix of all the systematics passing at least 10% correlation between one other systematic is shown in 9.10.

The following sections describe each of the systematic uncertainties considered in the analyses.

9.4.1 Luminosity

The uncertainty on the combined 2015+2016 integrated luminosity is 2.1%. It is derived, following a methodology similar to that detailed in Ref. [122], from a preliminary calibration of the luminosity scale using x-y beam-separation scans performed in August 2015 and May 2016. The uncertainties for the 2015 and 2016 data sets are partially correlated.

9.4.2 Reconstructed objects

9.4.2.1 Leptons

Uncertainties associated with leptons arise from the trigger, reconstruction, identification, and isolation, as well as the lepton momentum scale and resolution. The reconstruction, identification and isolation efficiency of electrons and muons, as well as the efficiency of the trigger used to record the events, differ slightly between data and simulation, which is compensated for by dedicated scale factor (SFs).

Efficiency SFs are derived using tag-and-probe techniques on $Z \to \ell^+\ell^-$ ($\ell = e, \mu$) data and simulated samples, and are applied to the simulation to correct for differences. The effect of these uncertainties is propagated as corrections to the event weight.

The total uncertainty on the efficiency SFs is $< 0.5\%$ for muons across the entire p_{T} spectrum [123] and for electrons with $p_{\mathrm{T}} > 30$ GeV, while it exceeds 1% for low p_{T} electrons [124].

Additional sources of uncertainty originate from the corrections applied to adjust the lepton momentum scale and resolution in the simulation to match those in data, measured using reconstructed distributions of the $Z \to \ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$ masses, as well as the measured E/p in $W \to e\nu$ events, where E and

Systematic uncertainty	Type	Components
Luminosity	N	1
Reconstructed Objects		
Electron trigger+reco+ID+isolation	SN	4
Electron energy scale+resolution	SN	2
Muon trigger+reco+ID+isolation	SN	10
Muon momentum scale+resolution+saggita	SN	5
Pileup modeling	SN	1
Jet vertex tagger	SN	1
Jet energy scale	SN	20
Jet energy resolution	SN	1
Double b-tagging efficiency	SN	12
Background and Signal Model		
$t\bar{t}$ cross section	N (free floating)	1
$t\bar{t}+ \geq 1c$: normalization	N	1
$t\bar{t}+ \geq 3b$: normalization	N	1
$t\bar{t}+\geq 1b$: NLO Shape	SN	9
$t\bar{t}+\geq 1c$: NLO Shape	SN	1
$t\bar{t}$ modeling: residual Radiation	SN	3
$t\bar{t}$ modeling: residual NLO generator	SN	3
$t\bar{t}$ modeling: residual parton shower+hadronization	SN	3
$t\bar{t}$ modeling: residual 5FS vs 4FS	SN	1
W+jets global normalization	N	1
W+jets relative normalization	N	4
W+jets scales and PDF	SN	7
W+jets generator	S	1
Z+jets global normalizations	N (free floating)	1
Z+jets heavy flavor	S	2
Single top cross section	N	1
Single top model	SN	2
Diboson normalization	N	1
$t\bar{t}V$ cross section	N	4
$t\bar{t}V$ modeling	SN	2
$t\bar{t}WW$ cross section	N	2
Signal cross section	N	1
Signal acceptance (generator)	N	1
Signal acceptance (PS)	N	1

Table 9.1: The list of systematic uncertainties considered. An "N" means that the uncertainty is taken as normalization-only for all processes and channels affected, whereas "SN" means that the uncertainty is taken on both shape and normalization. Some of the systematic uncertainties are split into several components for a more accurate treatment.

p are the electron energy and momentum, as measured by the calorimeter and the tracker respectively. To evaluate the effect of momentum scale uncertainties, the event selection is redone with the lepton momentum varied by $\pm 1\sigma$. For the momentum resolution uncertainties the event selection is redone with the lepton momentum smeared.

9.4.2.2 Jets

Uncertainties associated with jets in principle arise from the efficiency of jet reconstruction and identification based on the JVT variable, as well as the jet energy scale (JES) and resolution. Since reclustered jets are made by simply clustering the anti-k_t $R = 0.4$ jets, the uncertainties that get assigned the those jets will propagate to the reclustered jets.

The jet energy scale and its uncertainty were derived by combining information from test-beam data, LHC collision data and simulation [125]. The 50 nuisance parameters from the in situ analyses and the terms which form the extrapolation from 2012 to 2015 uncertainties have been combined to form 9 nuisance parameters (7 dominant eigenvectors, 1 residual term, 1 term for the extrapolation uncertainties, which have a strong η dependence). The remaining terms detailed above are combined as per recommendation via grouping of similar terms, and provide the other 11 nuisance parameters to form a total of 20. The JES uncertainty is about 5.5% for jets with $p_{\mathrm{T}} = 25$ GeV and quickly decreases with increasing jet p_{T}. It is below 1.5% for central jets with p_{T} in the range of $\simeq 100$ GeV–1.5 TeV. It represents one of the leading sources of uncertainty associated with reconstructed objects, affecting the relative normalizations of signal and the $t\bar{t}$ backgrounds in the different bins of number of jets.

The jet energy resolution was measured in Run 1 data and simulation as a function of the jet p_{T} and rapidity using dijet events. They were found to agree within 10% [126]. Additional uncertainties have been assessed in the extrapolation from Run 1 to Run 2 conditions [125]. The combined uncertainty is propagated by smearing the jet p_{T} in the simulation.

Additional small uncertainties associated with the modeling of the underlying event, in particular its impact on the p_{T} scale and resolution of unclustered energy, are also taken into account.

9.4.2.3 Double b-tagging

As mentioned in section refsub:results, p_T dependent scale factors were derived from g $\rightarrow$ bb events to correct MC double tagging efficiency for either (b, b) or (b, ℓ) reclustered jets. Four major sources of uncertainty contribute to affect the SF systematics. Two sources come from the varying the jet energy scale (1 variation for each component of the JES uncertainty) and the track smearing and measuring the direct impact on the SF. Another source comes from the fact that the dijet MC samples used to make the SF measurement had unreliable flavor fractions. A separate measurement to correct the dijet MC flavor fractions was performed, and the error of this measurement was propagated through to the SF measurement. The final source of uncertainty is due to the g $\rightarrow$ bb sample being biased in the tagging efficiency due to the event selection requiring the g $\rightarrow$ bb reclustered jet candidate having one muon ΔR matched to enrich bb purity. The full

SF measurement was repeated requiring 2 muons in the candidate jet and the average fractional difference between the nominal SF of the single and double muon sets was taken as an overall uncertainty to account for the bias.

To properly handle the correlations between the various sources of uncertainty across p_T bins and flavor, a 12×12 total covariance matrix was constructed by summing the statistical covariance matrix and the covariance matrix for each aforementioned systematic. Each element for a given systematic covariance matrix is computed with:

$$C_{ij} = \sum_{\alpha \in JES} (\text{Nom}_i - \text{JES}_i^{\alpha})(\text{Nom}_j - \text{JES}_j^{\alpha}) \tag{9.6}$$

where Nom_i is a nominally measured SF and JES_i^{α} is a SF measured after varying a given systematic. It is assumed that each covariance matrix element has a correlation of 1. For the muon bias uncertainty, the average fractional difference was applied to each SF and added to the diagonals of the matrix.

After getting the total covariance matrix, the matrix is diagonalized to yield 12 eigenvalues and eigenvectors. These result in 12 variations that can be applied to the SF when used in the analysis by applying:

$$\sqrt{(\lambda_m)}\text{Eigenvector}[m][\text{flavor}] \tag{9.7}$$

where m is between 1 through 12, λ_m is the m-th eigenvalue, and Eigenvector$[m]$[flavor] is the component of the m-th eigenvector associated to the SF based on flavor, p_T, and WP. Each of the 12 variations from all the eigenvectors were considered as separate systematics. Figures 9.7,9.8,9.9, show the statistical correlation matrix, the total correlation matrix, and the resulting eigenvectors used, respectively.

9.4.3 Background and signal modeling

9.4.3.1 $t\bar{t}$ +jets

A number of systematic uncertainties affecting the modeling of the $t\bar{t}$+jets background are considered. A free floating parameter is used for the inclusive $t\bar{t}$ NNLO+NNLL production cross-section [98], including effects from varying the factorization and renormalization scales, the PDF, α_S, and the top quark mass. Additional uncertainties related to the choice of NLO generator, PS and hadronization model, are assessed by comparing the nominal $t\bar{t}$ samples with alternative predictions produced by different MC generators. Any differences between samples predicted yields are taken as uncertainties. The normalization of $t\bar{t}$+$c\bar{c}$ is conservatively allowed to vary within a 50% in the fit as well.

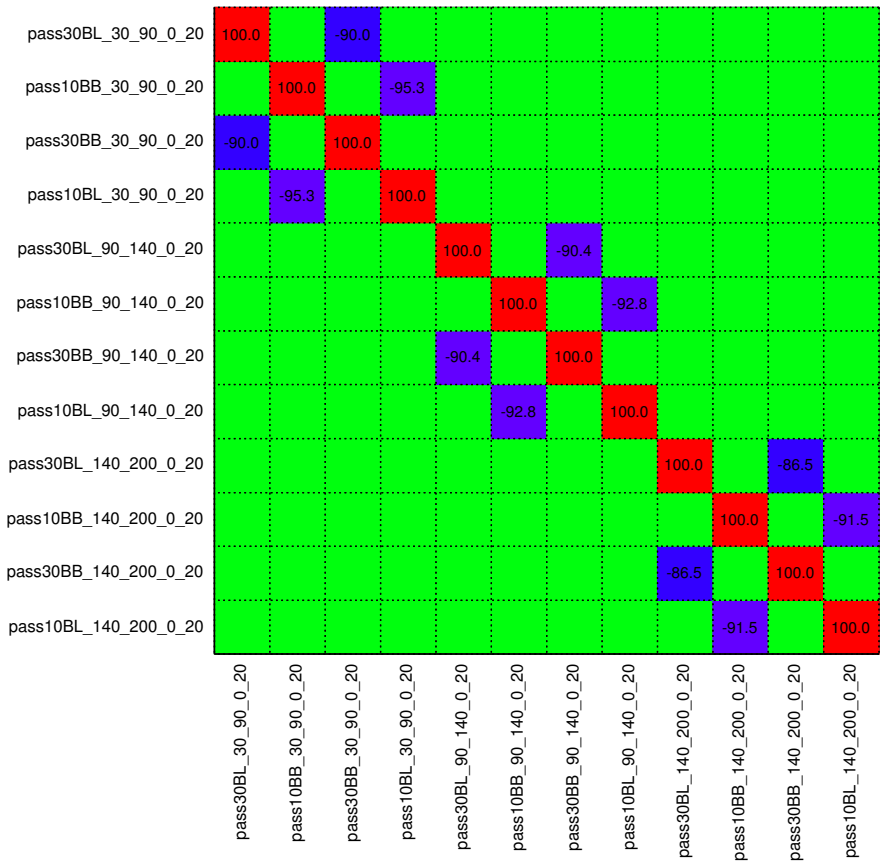

Figure 9.7: The statistical correlation between all components corresponding to the proper flavor, p_T bin, and WP as labeled.

9.4.3.2 Z+jets

For the normalization of Z+jets, the uncertainty on the global normalization is left as a free parameter in the profile-likelihood fit. The choice of signal and control regions is sensitive to the b-quark multiplicity in the Z+jets background events. Specifically, the use of the loose WP enriches the Z+jets sample with jets with 1 b-quark, while the tight WP enriches the Z+jets sample with jets with 2 b-quarks. Since both WP are used to define the signal and control regions, the contributions of Z+jets with different b-quark multiplicities will vary from region to region. In other words, the control region using two loose double tagged jets will typically have less total b-quarks compared to the signal regions using the tight WP to double tag jets. Thus, uncertainties were applied to events based on truth b-quark multiplicity in bins from 0 to 2 inclusively and greater than 2. Both samples split by b-quark multiplicity were each allowed to vary by 30% in the fit.

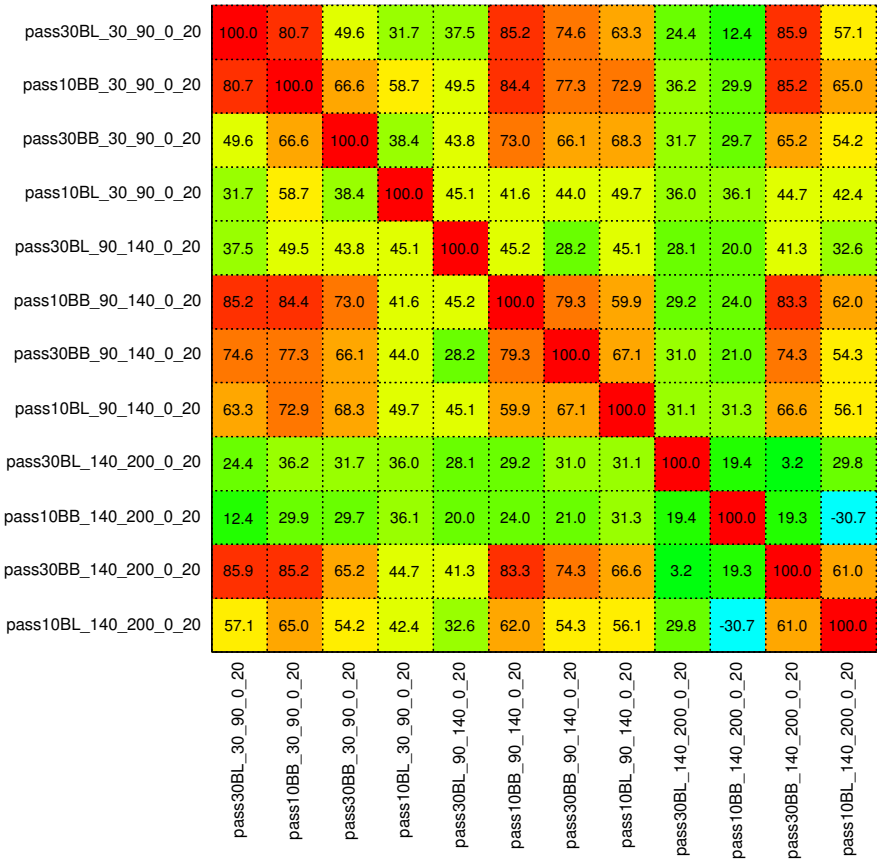

Figure 9.8: The total correlation between all components corresponding to the proper flavor, p_T bin, and WP as labeled.

9.4.3.3 Other simulated backgrounds

An uncertainty of $^{+5\%}_{-4\%}$ is used on the total cross-section for single-top production [119–121]. An additional uncertainty on initial- and final-state radiation is evaluated in a manner similar to that used for $t\bar{t}$. The uncertainty on the interference between Wt and $t\bar{t}$ production at NLO [127] is assessed by comparing the default "diagram removal" scheme to an alternative "diagram subtraction" scheme. A 50% normalization uncertainty on the diboson background is used, which includes uncertainties on the inclusive cross-section and additional jet production [128]. The uncertainty on the $t\bar{t}V$ NLO cross-section prediction is 15% [88]. A total 50% normalization uncertainty is used on the 4-top background.

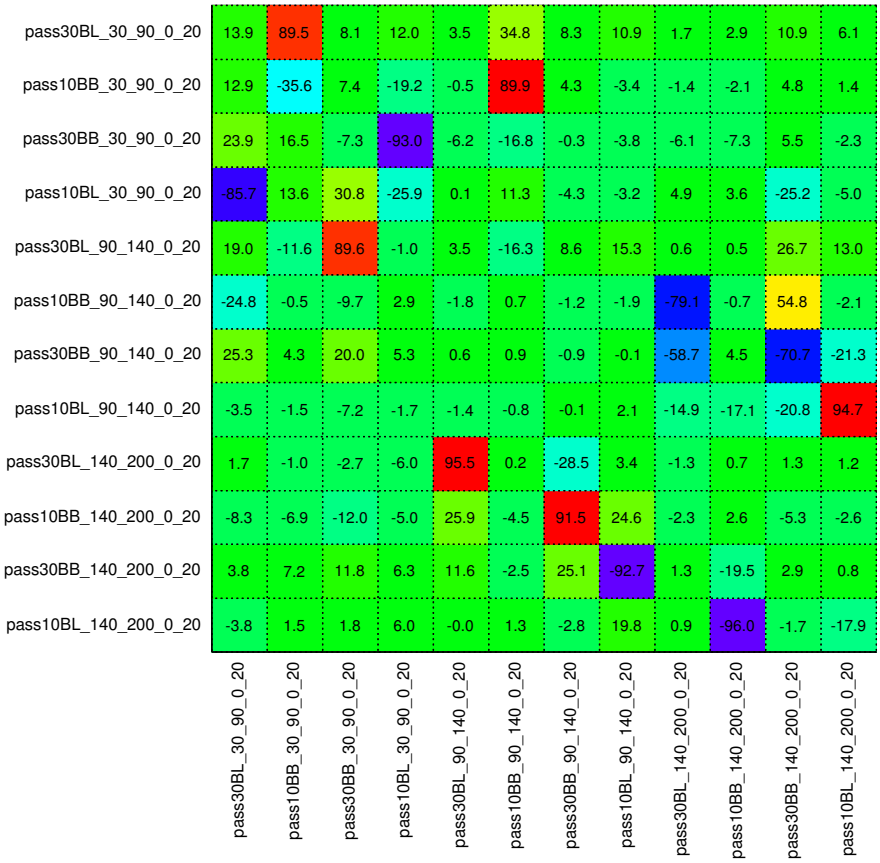

Figure 9.9: The resulting eigenvectors (bottom). Eigenvectors are reported as columns with each component corresponding to the proper flavor, p_T bin, and WP as labeled.

9.4.3.4 Signal

An uncertainty associated to the acceptance of the signal in the different regions due to different scale and PDF choices is also considered, and found to affect the normalization of the processes in the different regions by approximately $\pm 4\%$.

Finally, the effect on the acceptance due to the choice of parton shower is also accounted for, and found to be $\pm 6\%$ in all regions. In this case, the effect of using Herwig7 instead of the nominal Pythia8 has been parameterized as function of the p_{T} of the VH (where V is a W or Z) system in VH, $H \to b\bar{b}$ samples, and the effect propagated to the $H \to 2a \to 4b$ signal samples[1]. Figure 9.11 shows the parameterization of this uncertainty.

[1]The $H \to 2a \to 4b$ decay is not available in Herwig7.

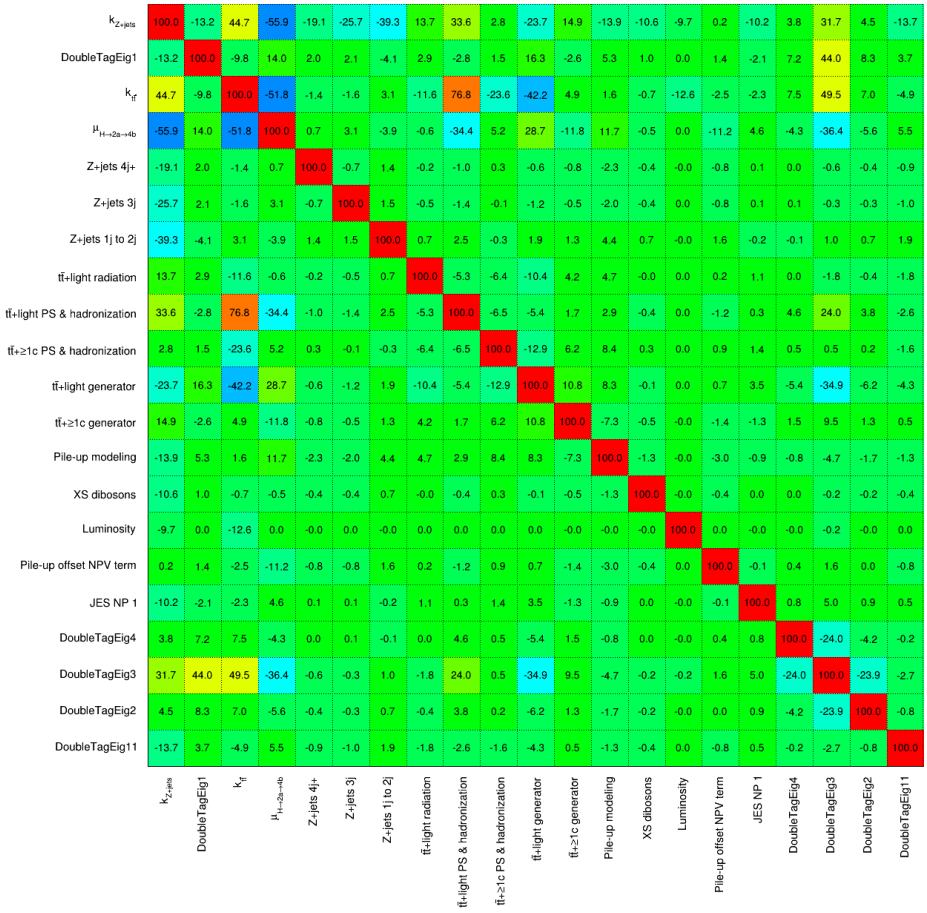

Figure 9.10: Correlation matrix between all systematics and normalization factors.

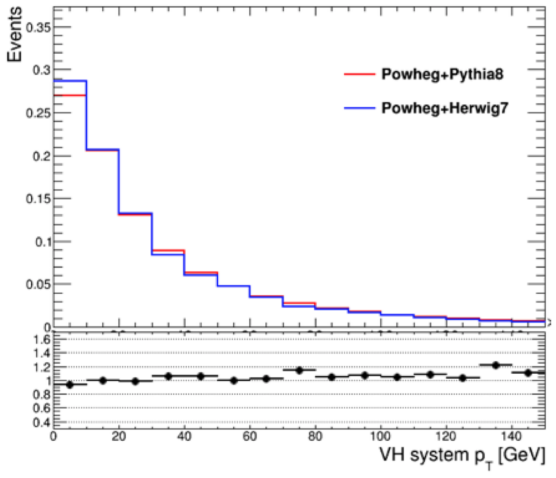

Figure 9.11: Comparison of the p_{T} spectra of the VH system when considering Pythia8 or Herwig7 parton showers.

9.5 Statistical analysis

Total yields in each of the regions considered (control and signal regions) are combined to test for the presence of a signal. The statistical analysis is based what is described in chapter 5. A likelihood function $\mathcal{L}(\mu, \theta)$ constructed as a product of Poisson probability terms over all bins defined from the signal and control regions described in section 9.3. This function depends on the signal-strength parameter μ, a multiplicative factor to the theoretical signal production cross section (for ZH production), and θ, a set of nuisance parameters that encode the effect of statistical or systematic uncertainties on the signal and background expectations and that are implemented in the likelihood function as log-normal or Gaussian priors, respectively.

The normalization of each background is determined from the fit simultaneously with μ, constrained by the uncertainties of the respective theoretical calculations, the uncertainty on the luminosity, and the data themselves. Statistical uncertainties in each bin of the discriminant distributions are taken into account by dedicated parameters in the fit.

Systematic uncertainties are pruned away and smoothed in order to reduce the CPU time for fitting and make the minimization more stable. The shape and the normalization of systematics are pruned separately per region and per sample. The pruning algorithm removes systematic uncertainties that have a $\leq$1% impact on normalization or those that cause a negligibly small impact on the shape of the final discriminant (all bins are within 1% of the nominal shape).

The test statistic used is the one described in section 5.4 and is used to measure the compatibility of the observed data with the signal plus background hypothesis (i.e. the exclusion test). For a given signal scenario, values of the branching fractions of $H \to 2a \to 4b$ (parameterized by μ) yielding $\mathrm{CL_s}<0.05$, where $\mathrm{CL_s}$ is computed using the asymptotic approximation [129, 130], are excluded at $\geq$95% CL.

Detailed validation studies of the fitting procedure have been performed using the simulation. To verify the improved background prediction, fits are performed under the background-only hypothesis. Statistical uncertainties in each bin of the discriminant distributions are also taken into account by dedicated parameters in the fit.

For the final result, a simultaneous fit to the data is performed on the distributions of the discriminants in 6 regions each with one bin. The fits are performed under a signal-plus-background hypothesis, where the signal-strength parameter μ is the parameter of interest in the fit, allowed to float freely but correlated through all regions.

9.6 Results

Chapter 5 describes how a binned likelihood fit is used to search for the signal in general, while section 9.5 outlines the specifics for the $H \to 2a \to 4b$ analysis. As explained in Sec. 9.3, the dilepton channel was examined by looking for double b-jets and investigating mass distributions of the tagged reclustered jets. The total yields in both the signal and control regions are fit.

The fit is performed separately for each signal considered, corresponding to a different mass for the new spin-zero particle, m_a. The results for the $m_a = 20$ GeV signal mass hypothesis are described in this section, while the fit results for the other mass hypotheses considered are included in Appendix A. The results for the various signal mass hypotheses are consistent with each other.

9.6.1 Fit inputs

The basic comparison of data and MC yields are checked in background-enriched regions (control and validation) with full systematic uncertainties. Figure 9.12 shows the data and MC yields in all the regions in the analysis except the signal regions. The signal regions are blinded, i.e. the data in the signal regions is not shown. The blinding consists of removing the bins in the signal region distributions with a signal-over-background larger than 1% when the signal is normalized to the SM cross section ZH production. This blinding has also been found to be independent of the mass of the a-boson.

The yield agreement in all unblinded control regions is reasonable and within expected errors. The large uncertainties are mostly attributed to the uncertainty on the applied double b-tagging SF.

9.6.2 Fits to the Asimov dataset

This section shows the signal+background fits to the Asimov dataset for the aforementioned regions. The fitted regions are shown in Figure 9.13 Figure 9.14 show the fitted nuisance parameters, normalization factors and gamma parameters to the Asimov dataset. The expected background is fitted to the Asimov dataset (the nuisance parameters are set at 0 while the normalization factors are set at 1, by definition) under the signal+background hypothesis.

Figure 9.14 shows the fitted nuisance parameters, normalization factors and gamma parameters to the Asimov dataset. Nearly all of the systematics remain unconstrained except for the $t\bar{t}$+light-jets generator and one double b-tagging eigenvector. The $t\bar{t}$+light-jets generator systematics are done by producing an alternative sample and comparing the differences between the nominal as described in section 9.4.3.1. These alternative samples are produced with less statistics than the nominal and as a result, very little statistics

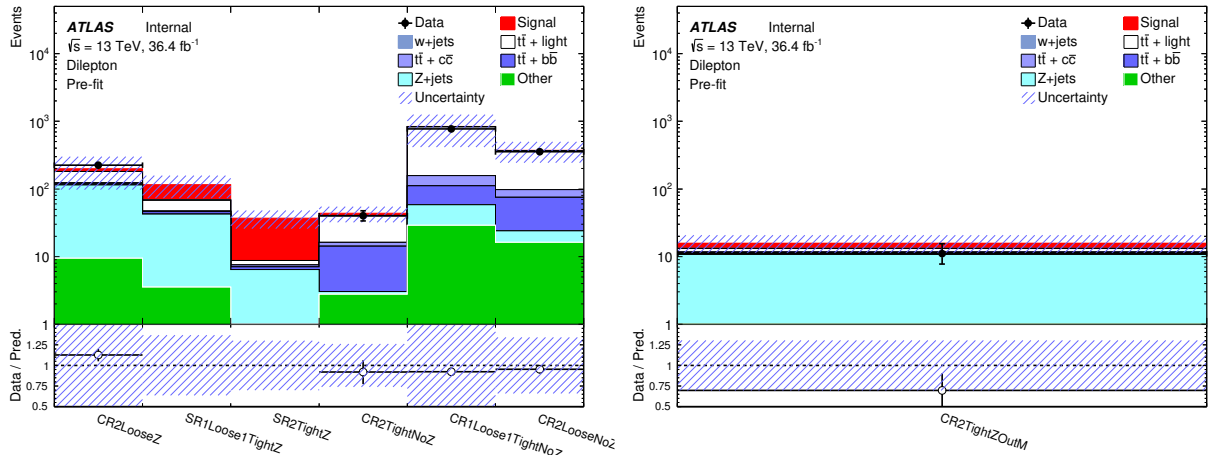

Figure 9.12: Signal and control regions (left) and validation regions (right) for the analysis showing the background composition and the expected signal yield for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$. The data and MC yields are compared in the control and validation regions, while the signal regions are blinded and no data is shown.

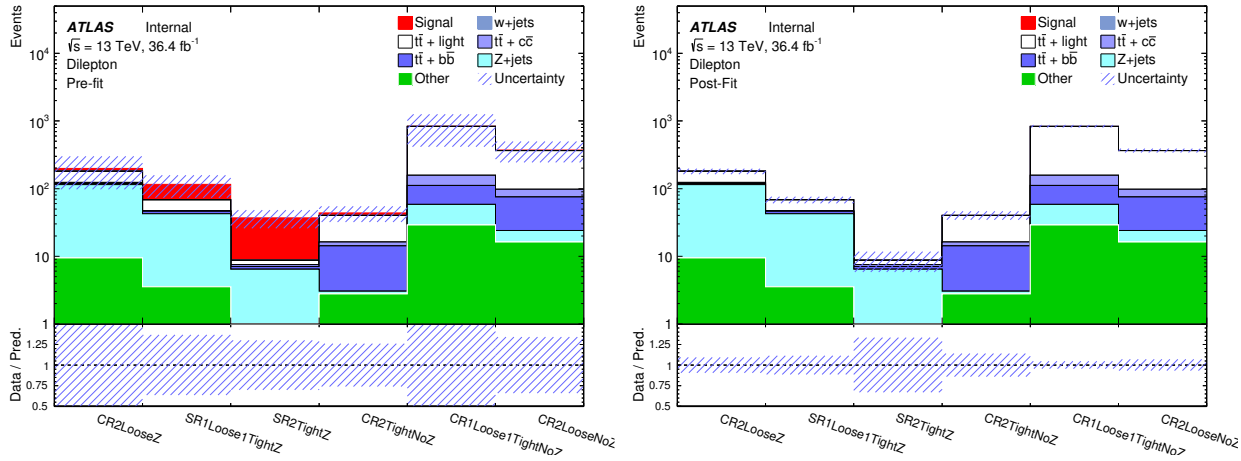

Figure 9.13: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

remain after requiring the selections for the signal and control regions. The statistical fluctuations cause an overestimation in the prior fed into the fit.

The third eigenvector is dominated by the tight WP (b, ℓ) flavor $90 - 140$ GeV bin component, which can be seen in Fig 9.9. This particular component is well represented in each region, providing sufficient statistics to have constraint.

Figure 9.14 also shows the effect of the different nuisance parameters on the parameter of interest. The

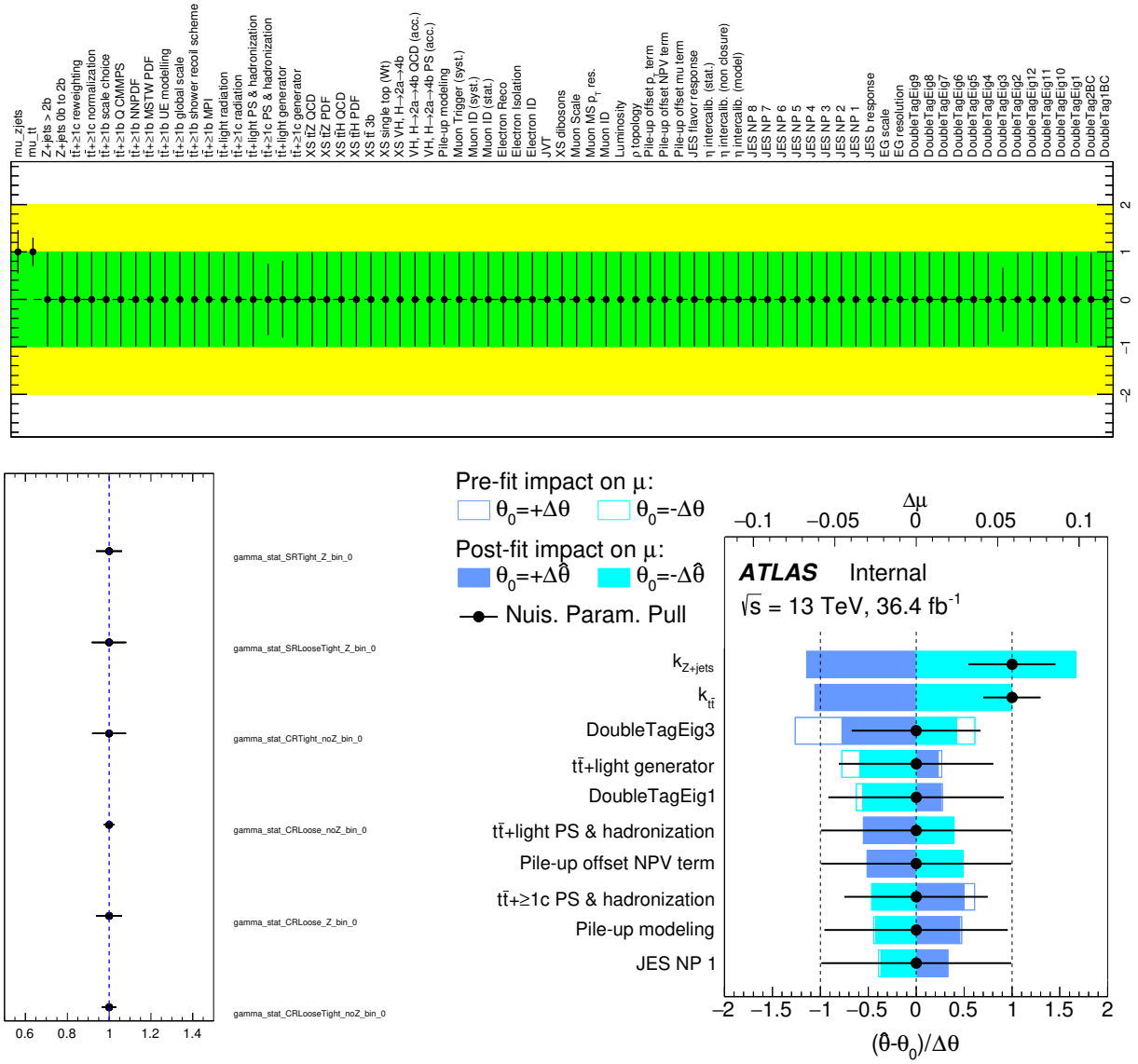

Figure 9.14: Nuisance parameters and normalization factors (top), gamma parameters (bottom left) and rankings based on the impact on the signal strength (bottom right) for the Asimov data fit using the full set of systematics.

parameters are ordered by post-fit $\Delta\mu$. Nuisance parameters related to the double b-tagging as well as the background normalization and modeling uncertainties have the largest effect on the signal strength.

9.6.3 Fits to blinded data

This section shows results for the analysis using full background-only fits to data with the blinded signal regions. Figure 9.15 shows a comparison of the yields in data and simulation for the control regions before and after the blinded fit to data. The yields from simulation in the signal regions are also shown. These plots show how the shaded area representing the total statistical and systematic uncertainties is reduced due to the profiling of the fit, and the improvement of the data/MC agreement.

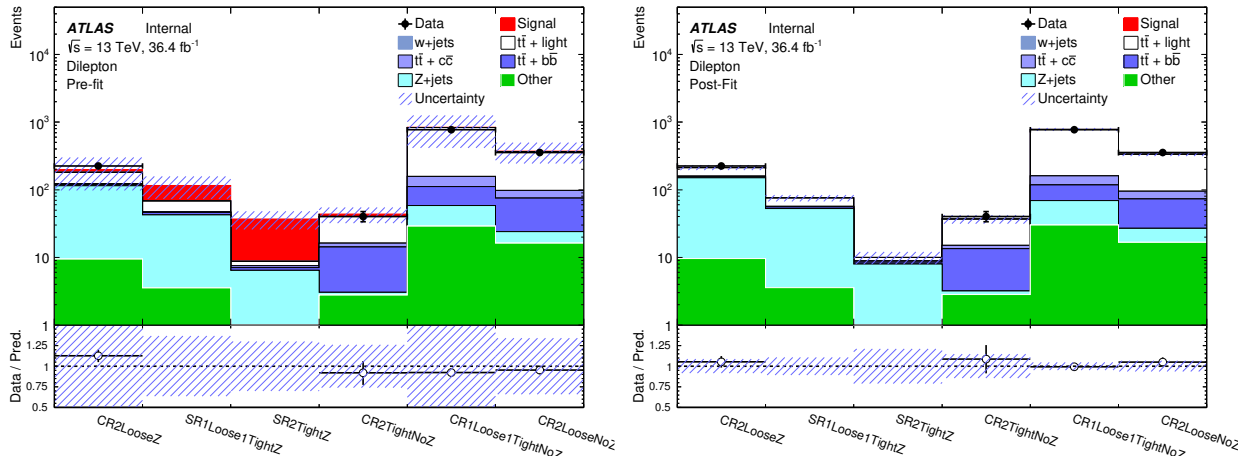

Figure 9.15: Predicted yields for the signal and control regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

The nuisance parameters, as well as the normalization factors and the gamma parameters for the background-only fit are shown in Figure 9.15. The constraints in the different nuisance parameters are consistent with those found in the Asimov fits, but since these fits are with real data, some pulls are observed. This is expected since the data is not in perfect agreement with the MC prediction. However, all pulled systematics have errors that comfortably cover the nominal values.

9.6.4 Fits to unblinded data

This section shows results for the analysis using full signal+background fits to data with the unblinded signal regions. Figure 9.17 shows a comparison of the yields in data and simulation for all regions before and after the unblinded fit to data. These plots show how the shaded area representing the total statistical and systematic uncertainties is reduced due to the profiling of the fit, and the improvement of the data/MC

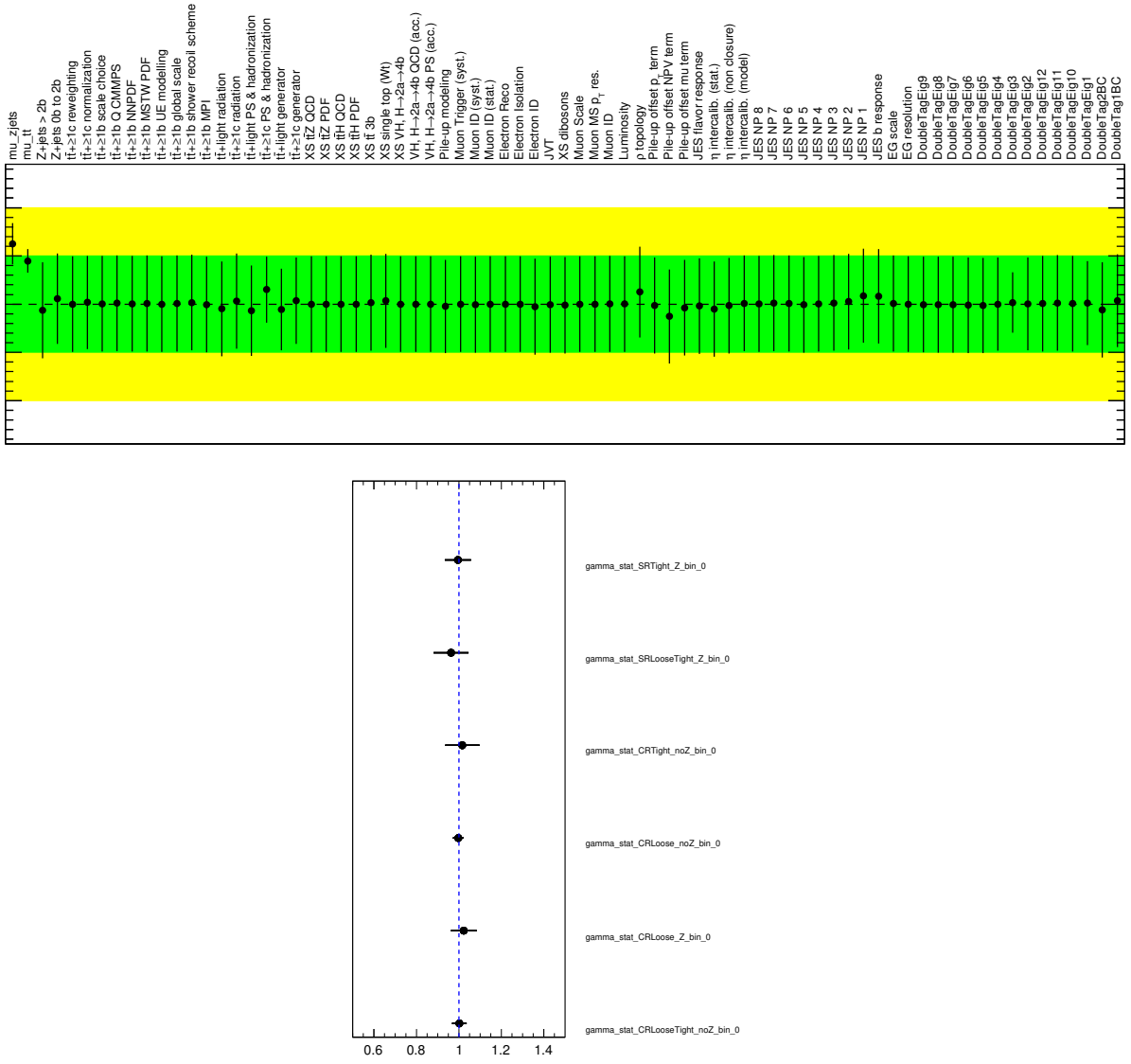

Figure 9.16: Nuisance parameters and normalization factors (top), gamma parameters (bottom) for the blinded data fit using the full set of systematics.

agreement.

The nuisance parameters, as well as the normalization factors and the gamma parameters for the signal+background fit are shown in Figure 9.17. The constraints in the different nuisance parameters are consistent with those found in the Asimov fits and background-only fits with data.

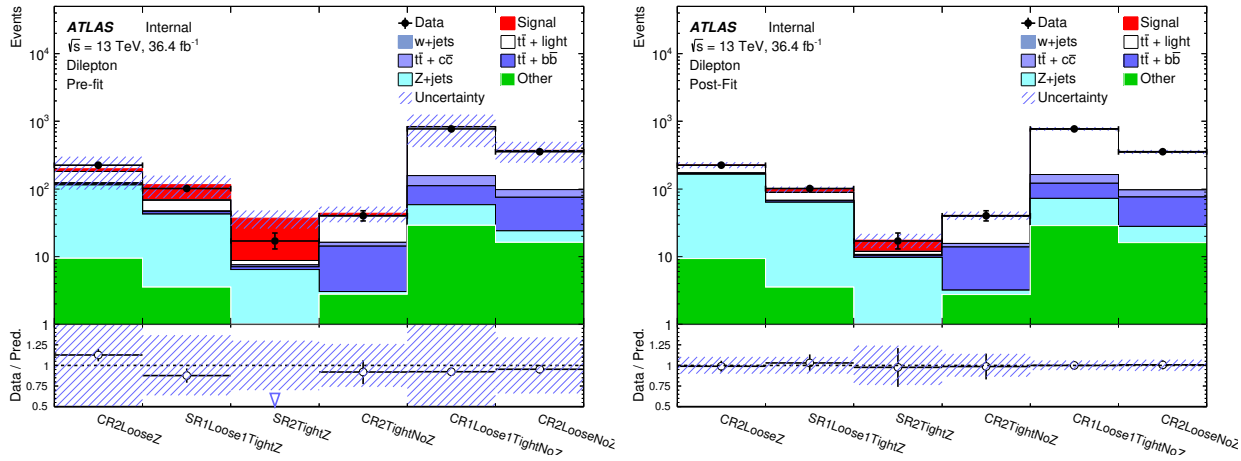

Figure 9.17: Predicted yields for the signal and control regions in the unblinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 20$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and $\mathrm{BR}(a \to bb) = 1$ is shown before the fit, while the signal yields set to the best fit signal strength in the fit (signal+background hypothesis).

9.6.5 Cross section limits on $H \to 2a \to 4b$ production

Prior to unblinding the analysis, the fit to the Asimov dataset can be used to derive expected upper limits on the $H \to 2a \to 4b$ production cross section times branching fraction. The results are shown in Figure 9.19 for the mass range $15 < m_a < 30$ GeV. The total signal cross section is parameterized in terms of the associated Higgs boson production cross section σ_{ZH},

$$\sigma_{4b} = C_{4b}^2 \times \sigma_{ZH}, \tag{9.8}$$

where

$$C_{4b}^2 = \kappa_{HZZ}^2 \times \mathrm{BR}(H \to 2a) \times \mathrm{BR}(a \to 2b)^2. \tag{9.9}$$

Figure 9.20 shows the previous resolved analysis results on the ZH production cross section times branching fraction again for ease of comparison. The previous results had an expected limit of **1.2 pb** and observed limit of **1.6 pb** at the 20 GeV mass point. The new strategy using the double b-tagger and the updated analysis strategy yields an expected limit of **0.34 pb**, improving the previous result by more than a factor of two. For reference, the cross-section of SM ZH production is **0.88 pb**. A comparison of the expected limits from the strategy using the double tagger and the previous results are shown in Fig 9.21. The mass points at 30 GeV and above are very similar.

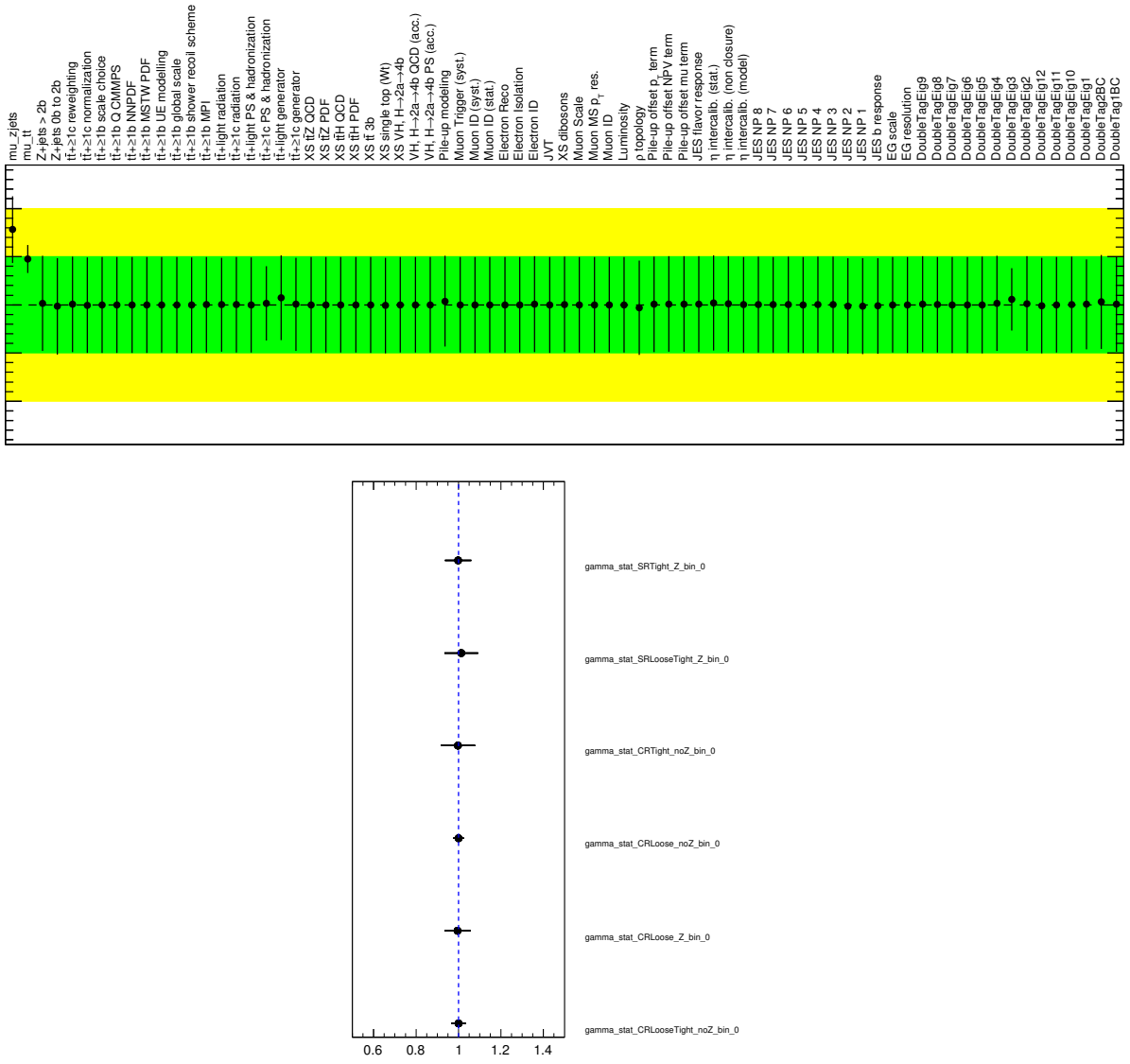

Figure 9.18: Nuisance parameters and normalization factors (top), gamma parameters (bottom) for the unblinded data fit using the full set of systematics.

The improved results indicate that the double b-tagger does indeed improve the sensitivity of the search for low mass signatures decaying into collimated b-quarks. Thus, a range of masses from $15 < m_a < 25$ GeV is covered by the merged search, while masses of 30 GeV and above are covered by the dedicated resolved analysis.

The observed result and limits can be seen and compared to the expected limits after unblinding. Fig-

ure 9.22 shows the observed limits for all the considered mass points after the unblinded fit. No significant excess of data events is observed in the signal regions and thus the measurement constrains the $H \to 2a \to 4b$ production cross section times branching fraction.

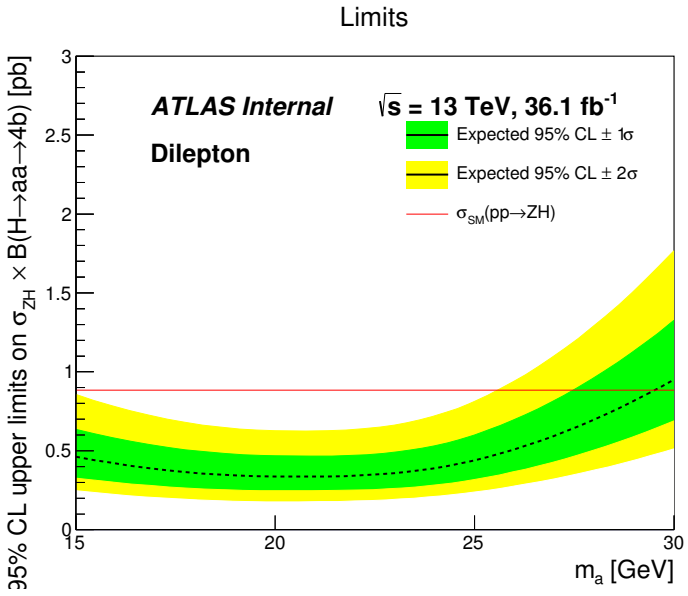

Figure 9.19: Expected 95% CL upper limits on the ZH, $H \to 2a \to 4b$ cross section times branching fraction as a function of the a mass for the mass range $15 < m_a < 30$ GeV. The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction and its ± 1 standard deviation uncertainty.

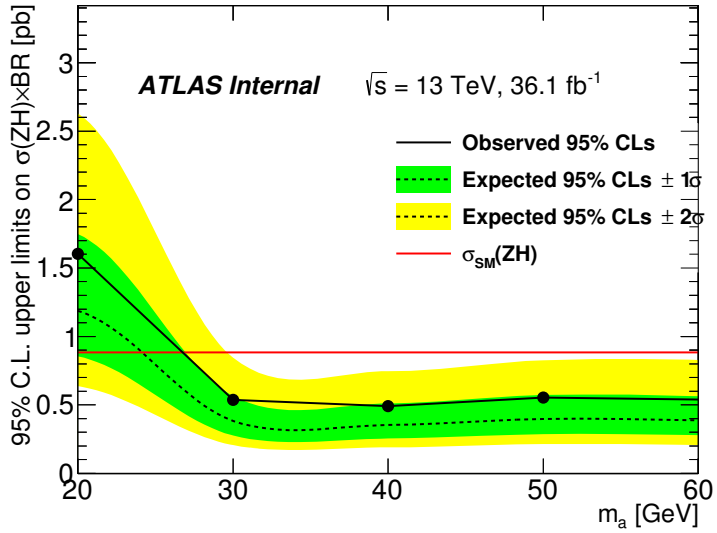

Figure 9.20: Interpretation results from previous resolved analysis [131, 132]. Expected 95% CL upper limits on the ZH, $H \to 2a \to 4b$ cross section times branching fraction as a function of the a mass for the mass range $20 < m_a < 60$ GeV. The surrounding shaded bands correspond to the ± 1 and ± 2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction and its ± 1 standard deviation uncertainty.

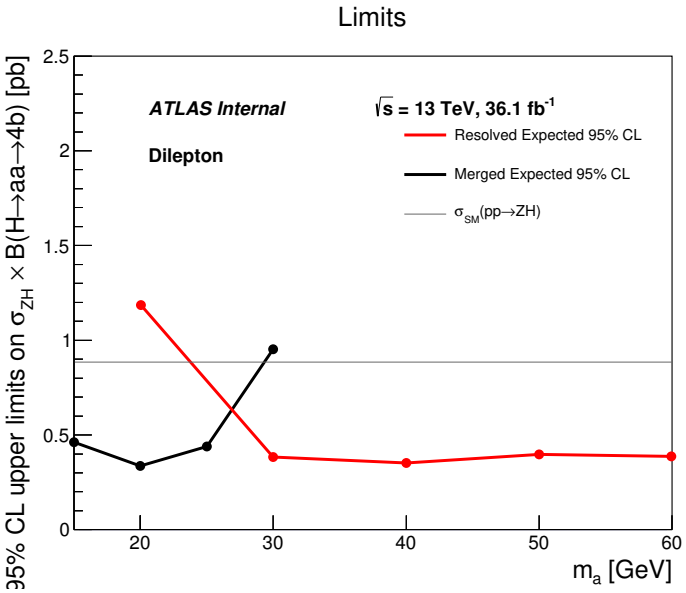

Figure 9.21: Comparison of the expected limits for this merged result and the previous resolved analysis. These correspond to the expected 95% CL upper limits on the ZH, $H \to 2a \to 4b$ cross section times branching fraction as a function of the a mass for the mass range $15 < m_a < 60$ GeV.

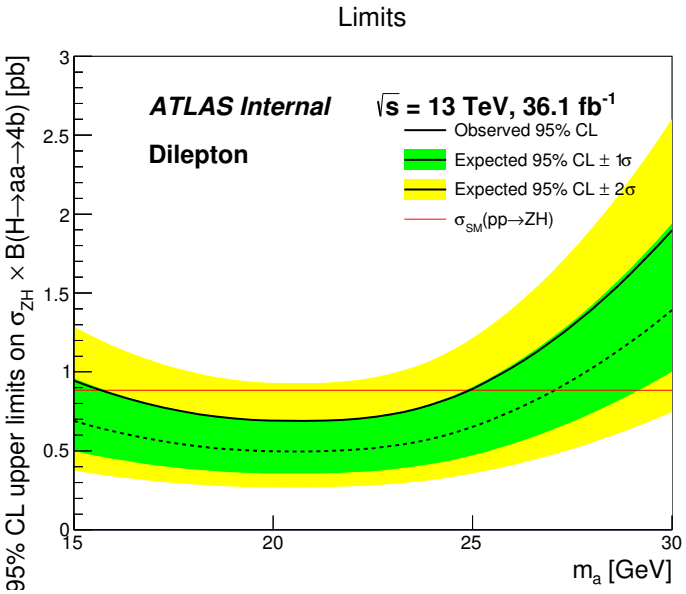

Figure 9.22: Observed limits for all considered mass points with the expected 95% CL upper limits on the ZH, $H \to 2a \to 4b$ cross section times branching fraction as a function of the a mass for the mass range $15 < m_a < 30$ GeV.

Chapter 10

Summary and conclusion

This thesis presented a search for $H \to 2a \to 4b$ in the low m_a regime using the pp data set at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector during 2015 and 2016. The Higgs production mode investigated was associated production with a Z boson to have leptons to trigger on. At low m_a, the decaying b-quarks become collimated and traditional ATLAS reconstruction techniques are not sensitive to the signature. Thus, the major challenge for this analysis was developing a strategy to identify the $a \to bb$ resonances.

The development of a new, dedicated low mass double b-tagger designed to address this challenge was presented. A sample of truth labeled jets from MC samples of $m_a = 20$ GeV $H \to 2a \to 4b$ and $t\bar{t}$ processes was used to train a BDT classifier to identify low p_T large radius, reclustered jets that carry two b-hadrons. After performing an exhaustive search for potential discriminatory variables, the tagger was chosen to use variables derived from track subjets created from the exclusive-k_T algorithm. The subjet kinematics and b-tagging variables were used as inputs to the BDT.

After sufficient optimization of the BDT classifier, work on the calibration was done. The efficiency for reclustered jets with (b, b) and (b, ℓ) flavors was investigated and compared between data and MC. Scale factors and full consideration of their uncertainties were derived to correct the MC to behave like data.

These scale factors were then used in the low mass $H \to 2a \to 4b$ search. Signal and control regions were defined by finding events with two leptons and two double b-tagged reclustered jets. The WP used for tagging varied depending on if the signal or control region was being defined since the looser definition allows for fakes to enrich backgrounds for modeling. The mass was reconstructed for the lepton pair and the tagged reclustered jets to reconstruct the Z mass, a-boson mass, and Higgs mass, respectively. Mass cuts were applied to identify these objects.

After defining all the signal and control regions, the yields were used as inputs into a profile likelihood ratio. The profile likelihood ratio was fit and calibrate the backgrounds and measure the signal strength while profiling the sytematics. The major systematics considered come from the Z+jets MC sample modeling, $t\bar{t}$ MC sample modeling, and uncertainty from the double b-tagging scale factor measurement. After performing a fit using an Asimov dataset, expected limits were found to improve the result from the previous resolved

analysis by more than a factor of two, thus indicating that the double tagging analysis strategy is working as designed. The observed limits were also obtained after performing an unblinded fit to all regions. A small one σ excess of data events was found in the dominant signal region causing the observed limits to be roughly one σ above the expected limits. This is not considered a significant excess for discovery of a new signal and so the measurement constrains the $H \to 2a \to 4b$ production cross section times branching fraction.

Bibliography

[1] G. Arnison et al., *Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 GeV*, Physics Letters B **122**.1 (1983) 103, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/0370269383911772.

[2] *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: arXiv:1207.7214[hep-ex] [hep-ex].

[3] *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: arXiv:1207.7235[hep-ex] [hep-ex].

[4] *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment* (2015), arXiv: arXiv:1507.04548[hep-ex] [hep-ex].

[5] *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, Eur. Phys. J. **C75**.5 (2015) 212, arXiv: arXiv:1412.8662[hep-ex] [hep-ex].

[6] G. Aad et al., *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV* (2016), arXiv: 1606.02266 [hep-ex].

[7] ATLAS Collaboration, "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment," tech. rep. ATLAS-CONF-2018-031, CERN, July 2018, URL: https://cds.cern.ch/record/2629412.

[8] S. Dawson et al., *Working Group Report: Higgs Boson* (2013), arXiv: arXiv:1310.8361[hep-ex] [hep-ex].

[9] D. Curtin et al., *Exotic decays of the 125 GeV Higgs boson*, Phys. Rev. **D90**.7 (2014) 075004, arXiv: arXiv:1312.4992[hep-ph] [hep-ph].

[10] B. A. Dobrescu and K. T. Matchev, *Light axion within the next-to-minimal supersymmetric standard model*, JHEP **09** (2000) 031, arXiv: hep-ph/0008192 [hep-ph].

[11] U. Ellwanger et al., *Towards a no lose theorem for NMSSM Higgs discovery at the LHC* (2003), arXiv: hep-ph/0305109 [hep-ph].

[12] R. Dermisek and J. F. Gunion, *Escaping the large fine tuning and little hierarchy problems in the next to minimal supersymmetric model and h —¿ aa decays*, Phys. Rev. Lett. **95** (2005) 041801, arXiv: hep-ph/0502105 [hep-ph].

[13] S. Chang et al., *Nonstandard Higgs Boson Decays*, Ann. Rev. Nucl. Part. Sci. **58** (2008) 75, arXiv: 0801.4554 [hep-ph].

[14] D. E. Morrissey and A. Pierce, *Modified Higgs Boson Phenomenology from Gauge or Gaugino Mediation in the NMSSM*, Phys. Rev. **D78** (2008) 075029, arXiv: 0807.2259 [hep-ph].

[15] V. Silveira and A. Zee, *SCALAR PHANTOMS*, Phys. Lett. **161B** (1985) 136.

[16] M. Pospelov, A. Ritz, and M. B. Voloshin, *Secluded WIMP Dark Matter*, Phys. Lett. **B662** (2008) 53, arXiv: arXiv:0711.4866[hep-ph] [hep-ph].

[17] P. Draper et al., *Dark Light Higgs*, Phys. Rev. Lett. **106** (2011) 121805, arXiv: arXiv:1009.3963[hep-ph] [hep-ph].

[18] S. Ipek, D. McKeen, and A. E. Nelson, *A Renormalizable Model for the Galactic Center Gamma Ray Excess from Dark Matter Annihilation*, Phys. Rev. **D90**.5 (2014) 055021, arXiv: arXiv:1404.3716[hep-ph] [hep-ph].

[19] A. Martin, J. Shelton, and J. Unwin, *Fitting the Galactic Center Gamma-Ray Excess with Cascade Annihilations*, Phys. Rev. **D90**.10 (2014) 103513, arXiv: arXiv:1405.0272[hep-ph] [hep-ph].

[20] S. Profumo, M. J. Ramsey-Musolf, and G. Shaughnessy, *Singlet Higgs phenomenology and the electroweak phase transition*, JHEP **08** (2007) 010, arXiv: arXiv:0705.2425[hep-ph] [hep-ph].

[21] N. Blinov et al., *Electroweak Baryogenesis from Exotic Electroweak Symmetry Breaking* (2015), arXiv: arXiv:1504.05195[hep-ph] [hep-ph].

[22] G. Burdman et al., *Folded supersymmetry and the LEP paradox*, JHEP **02** (2007) 009, arXiv: hep-ph/0609152 [hep-ph].

[23] N. Craig et al., *Naturalness in the Dark at the LHC*, JHEP **07** (2015) 105, arXiv: arXiv:1501.05310[hep-ph] [hep-ph].

[24] D. Curtin and C. B. Verhaaren, *Discovering Uncolored Naturalness in Exotic Higgs Decays* (2015), arXiv: arXiv:1506.06141[hep-ph] [hep-ph].

[25] L. Goodenough and D. Hooper, *Possible Evidence For Dark Matter Annihilation In The Inner Milky Way From The Fermi Gamma Ray Space Telescope* (2009), arXiv: arXiv:0910.2998[hep-ph] [hep-ph].

[26] D. Hooper and L. Goodenough, *Dark Matter Annihilation in The Galactic Center As Seen by the Fermi Gamma Ray Space Telescope*, Phys. Lett. **B697** (2011) 412, arXiv: arXiv:1010.2752[hep-ph] [hep-ph].

[27] A. Boyarsky, D. Malyshev, and O. Ruchayskiy, *A comment on the emission from the Galactic Center as seen by the Fermi telescope*, Phys. Lett. **B705** (2011) 165, arXiv: arXiv:1012.5839[hep-ph] [hep-ph].

[28] D. Hooper and T. Linden, *On The Origin Of The Gamma Rays From The Galactic Center*, Phys. Rev. **D84** (2011) 123005, arXiv: arXiv:1110.0006[astro-ph.HE] [astro-ph.HE].

[29] K. N. Abazajian and M. Kaplinghat, *Detection of a Gamma-Ray Source in the Galactic Center Consistent with Extended Emission from Dark Matter Annihilation and Concentrated Astrophysical Emission*, Phys. Rev. **D86** (2012) 083511, [Erratum: Phys. Rev. **D87**, 129902 (2013)], arXiv: arXiv:1207.6047[astro-ph.HE] [astro-ph.HE].

[30] C. Gordon and O. Macias, *Dark Matter and Pulsar Model Constraints from Galactic Center Fermi-LAT Gamma Ray Observations*, Phys. Rev. **D88**.8 (2013) 083521, [Erratum: Phys. Rev. **D89**, no.4, 049901 (2014)], arXiv: arXiv:1306.5725[astro-ph.HE] [astro-ph.HE].

[31] K. N. Abazajian et al., *Astrophysical and Dark Matter Interpretations of Extended Gamma-Ray Emission from the Galactic Center*, Phys. Rev. **D90**.2 (2014) 023526, arXiv: arXiv:1402.4090[astro-ph.HE] [astro-ph.HE].

[32] T. Daylan et al., *The Characterization of the Gamma-Ray Signal from the Central Milky Way: A Compelling Case for Annihilating Dark Matter* (2014), arXiv: arXiv:1402.6703[astro-ph.HE] [astro-ph.HE].

[33] B. Zhou et al., *GeV excess in the Milky Way: The role of diffuse galactic gamma-ray emission templates*, Phys. Rev. **D91**.12 (2015) 123010, arXiv: arXiv:1406.6948[astro-ph.HE] [astro-ph.HE].

[34] *Fermi-LAT Observations of High-Energy Gamma-Ray Emission Toward the Galactic Center* (2015), arXiv: arXiv:1511.02938[astro-ph.HE] [astro-ph.HE].

[35] M. Ackermann et al., *Observations of M31 and M33 with the Fermi Large Area Telescope: A Galactic Center Excess in Andromeda?* Astrophys. J. **836**.2 (2017) 208, arXiv: 1702.08602 [astro-ph.HE].

[36] C. Boehm et al., *Extended gamma-ray emission from Coy Dark Matter*, JCAP **1405** (2014) 009, arXiv: arXiv:1401.6458[hep-ph] [hep-ph].

[37] C. Cheung et al., *NMSSM Interpretation of the Galactic Center Excess*, Phys. Rev. **D90**.7 (2014) 075011, arXiv: arXiv:1406.6372[hep-ph] [hep-ph].

[38] J. Guo et al., *NMSSM explanations of the Galactic center gamma ray excess and promising LHC searches*, Phys. Rev. **D91**.9 (2015) 095003, arXiv: arXiv:1409.7864[hep-ph] [hep-ph].

[39] A. Berlin, D. Hooper, and S. D. McDermott, *Simplified Dark Matter Models for the Galactic Center Gamma-Ray Excess*, Phys. Rev. **D89**.11 (2014) 115022, arXiv: arXiv:1404.0022[hep-ph] [hep-ph].

[40] R. Bartels, S. Krishnamurthy, and C. Weniger, *Strong support for the millisecond pulsar origin of the Galactic center GeV excess* (2015), arXiv: arXiv:1506.05104[astro-ph.HE] [astro-ph.HE].

[41] S. K. Lee et al., *Evidence for Unresolved Gamma-Ray Point Sources in the Inner Galaxy* (2015), arXiv: arXiv:1506.05124[astro-ph.HE] [astro-ph.HE].

[42] M. Quiros, "Finite temperature field theory and phase transitions," *High energy physics and cosmology. Proceedings, Summer School, Trieste, Italy*, 1999 187, arXiv: arXiv:9901312[hep-ph] [hep-ph].

[43] D. E. Morrissey and M. J. Ramsey-Musolf, *Electroweak baryogenesis*, New J. Phys. **14** (2012) 125003, arXiv: arXiv:1206.2942[hep-ph] [hep-ph].

[44] S. Profumo et al., *Singlet-catalyzed electroweak phase transitions and precision Higgs boson studies*, Phys. Rev. **D91**.3 (2015) 035018, arXiv: arXiv:1407.5342[hep-ph] [hep-ph].

[45] D. Curtin, P. Meade, and C.-T. Yu, *Testing Electroweak Baryogenesis with Future Colliders*, JHEP **11** (2014) 127, arXiv: arXiv:1409.0005[hep-ph] [hep-ph].

[46] M. Carena et al., *Higgs Signal for $H \to aa$ at Hadron Colliders*, JHEP **04** (2008) 092, arXiv: arXiv:0712.2466[hep-ph] [hep-ph].

[47] K. Cheung, J. Song, and Q.-S. Yan, *Role of $H \to \eta\eta$ in Intermediate-Mass Higgs Boson Searches at the Large Hadron Collider*, Phys. Rev. Lett. **99** (2007) 031801, arXiv: arXiv:0703149[hep-ph] [hep-ph].

[48] D. E. Kaplan and M. McEvoy, *Associated Production of Non-Standard Higgs Bosons at the LHC*, Phys. Rev. **D83** (2011) 115004, arXiv: arXiv:1102.0704[hep-ph] [hep-ph].

[49] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3**.08 (Aug. 2008) S08001, URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

[50] *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[51] T. Kawamoto et al., "New Small Wheel Technical Design Report," tech. rep. CERN-LHCC-2013-006. ATLAS-TDR-020, CERN, June 2013, eprint: https://cds.cern.ch/record/1552862.

[52] S. K.-w. Chan, D. Lopez Mateos, and J. Huth, "Micromegas Trigger Processor Algorithm Performance in Nominal, Misaligned, and Corrected Misalignment Conditions," tech. rep. ATL-COM-UPGRADE-2015-033, CERN, Dec. 2015, URL: https://cds.cern.ch/record/2113121.

[53] B. Andersson et al., *Parton fragmentation and string dynamics*, Physics Reports **97**.2 (1983) 31, ISSN: 0370-1573, URL: http://www.sciencedirect.com/science/article/pii/0370157383900807.

[54] T. Sjöstrand, *Jet fragmentation of multiparton configurations in a string framework*, Nuclear Physics B **248**.2 (1984) 469, ISSN: 0550-3213, URL: http://www.sciencedirect.com/science/article/pii/0550321384906072.

[55] B. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nuclear Physics B **238**.3 (1984) 492, ISSN: 0550-3213, URL: http://www.sciencedirect.com/science/article/pii/055032138490333X.

[56] G. Marchesini and B. Webber,
Monte Carlo simulation of general hard processes with coherent QCD radiation,
Nuclear Physics B **310**.3 (1988) 461, ISSN: 0550-3213,
URL: http://www.sciencedirect.com/science/article/pii/0550321388900892.

[57] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*,
Nucl. Instrum. Meth. A **506** (2003) 250.

[58] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*,
The European Physical Journal C **71**.2 (Feb. 2011) 1554, ISSN: 1434-6052,
URL: https://doi.org/10.1140/epjc/s10052-011-1554-0.

[59] T. Junk, *Confidence level computation for combining searches with small statistics*,
Nucl. Instrum. Meth. A **434** (1999) 435, arXiv: hep-ex/9902006.

[60] S. Frixione, P. Nason, and G. Ridolfi,
A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction,
JHEP **0709** (2007) 126, arXiv: 0707.3088 [hep-ph].

[61] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*,
JHEP **0411** (2004) 040, arXiv: hep-ph/0409146.

[62] S. Frixione, P. Nason, and C. Oleari,
Matching NLO QCD computations with Parton Shower simulations: the POWHEG method,
JHEP **0711** (2007) 070, arXiv: 0709.2092 [hep-ph].

[63] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv: 1002.2581 [hep-ph].

[64] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024,
arXiv: 1007.2241 [hep-ph].

[65] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*,
Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[66] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **09** (2014) 145,
arXiv: 1406.3660 [hep-ex].

[67] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823,
arXiv: 1005.4568 [physics.ins-det].

[68] J. Pumplin et al.,
New generation of parton distributions with uncertainties from global QCD analysis,
JHEP **07** (2002) 012, arXiv: hep-ph/0201195 [hep-ph].

[69] M. Aaboud et al., *Search for the Higgs boson produced in association with a vector boson and decaying into two spin-zero particles in the $H \to aa \to 4b$ channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **10** (2018) 031, arXiv: 1806.07355 [hep-ex].

[70] *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. C **72** (2012) 1909, arXiv: 1110.3174 [hep-ex].

[71] *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data* (2014), arXiv: 1404.2240 [hep-ex].

[72] *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*,
Eur. Phys. J. C **74** (2014) 3071, arXiv: 1407.5063 [hep-ex].

[73] M. Aaboud et al., *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, The European Physical Journal C **77**.3 (Mar. 2017) 195,
ISSN: 1434-6052, URL: https://doi.org/10.1140/epjc/s10052-017-4756-2.

[74] "Lepton isolation recommendations,"
https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/IsolationSelectionTool.

[75] ATLAS Collaboration, *Muon Combined Performance in Run 2 (25 ns runs)*,
ATL-COM-MUON-2015-093 (2015), URL: https://cds.cern.ch/record/2105495.

[76] W. Lampl et al., "Calorimeter Clustering Algorithms: Description and Performance,"
tech. rep. ATL-LARG-PUB-2008-002, CERN, Apr. 2008.

[77] M. Cacciari, G. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063,
arXiv: 0802.1189 [hep-ph].

[78] T. A. collaboration, *Tagging and suppression of pileup jets* (2014).

[79] ATLAS Collaboration,
Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector,
ATLAS-CONF-2015-029 (2015), URL: http://cdsweb.cern.ch/record/2037702.

[80] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector* (2015), arXiv: 1510.03823 [hep-ex].

[81] A. Collaboration, *Performance of b -jet identification in the ATLAS experiment*, Journal of Instrumentation **11**.04 (2016) P04008, URL: http://stacks.iop.org/1748-0221/11/i=04/a=P04008.

[82] "Impact of Alternative Inputs and Grooming Methods on Large-R Jet Reconstruction in ATLAS," tech. rep. ATL-PHYS-PUB-2017-020, CERN, Dec. 2017, URL: https://cds.cern.ch/record/2297485.

[83] "Jet reclustering and close-by effects in ATLAS run II," tech. rep. ATLAS-CONF-2017-062, CERN, July 2017, URL: https://cds.cern.ch/record/2275649.

[84] B. Nachman et al., *Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, JHEP **02** (2015) 075, arXiv: 1407.2922 [hep-ph].

[85] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Physics Letters B **659**.1 (2008) 119, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/S0370269307011094.

[86] M. Cacciari, G. P. Salam, and G. Soyez, *The catchment area of jets*, Journal of High Energy Physics **2008**.04 (2008) 005, URL: http://stacks.iop.org/1126-6708/2008/i=04/a=005.

[87] "Variable Radius, Exclusive-k$_T$, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS," tech. rep. ATL-PHYS-PUB-2017-010, CERN, June 2017, URL: https://cds.cern.ch/record/2268678.

[88] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, JHEP **07** (2012) 052, arXiv: 1204.5678 [hep-ph].

[89] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021 (2014), https://cds.cern.ch/record/1966419.

[90] G. Aad et al., *Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73**.8 (2013) 2518, arXiv: 1302.4393 [hep-ex].

[91] T. Sjöstrand, S. Mrenna, and P. Skands, *A Brief Introduction to Pythia 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[92] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph].

[93] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, arXiv e-prints, physics/0703039 (Mar. 2007) physics/0703039, arXiv: physics/0703039 [physics.data-an].

[94] C. Issever et al., "X→bb Tagging rel. 20.7: g→bb Supporting Document," tech. rep. ATL-COM-PHYS-2017-1064, CERN, July 2017, URL: https://cds.cern.ch/record/2275237.

[95] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph].

[96] ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes* (2011), ATL-PHYS-PUB-2014-021, URL: http://cdsweb.cern.ch/record/1966419.

[97] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[98] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].

[99] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. **B710** (2012) 612, arXiv: 1111.5869 [hep-ph].

[100] P. Bärnreuther, M. Czakon, and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t}$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv: 1204.5201 [hep-ph].

[101] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **1212** (2012) 054, arXiv: 1207.0236 [hep-ph].

[102] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the quark-gluon reaction*, JHEP **1301** (2013) 080, arXiv: 1210.6832 [hep-ph].

[103] M. Czakon, P. Fiedler, and A. Mitov, *The total top quark pair production cross-section at hadron colliders through $\mathcal{O}(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph].

[104] A. Mitov, "An update on top pair differential distributions at NNLO," URL: %5Curl%7Bhttps://indico.cern.ch/event/472719/contributions/2166640/attachments/1274235/1889797/AlexanderMitov-topwg-2016.pdf%7D.

[105] D. H. R. de Moriond QCD 2016, "Differential top-quark pair production at NNLO," URL: %5Curl%7Bhttp://moriond.in2p3.fr/QCD/2016/SundayAfternoon/Heymes.pdf%7D.

[106] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a bb pair in pp collisions at sqrt s = 13 TeV with the ATLAS detector*, ATLAS-CONF-2016-080 (2016), URL: http://cdsweb.cern.ch/record/2206255.

[107] C. Helsens et al.,
Search for heavy vector-like quarks decaying to a Higgs boson and a third generation quark in the lepton plus jets final state in pp collisions at $\sqrt{s} = 8$ TeV *with the ATLAS detector*, ATL-COM-PHYS-2014-337 (2014),

[108] F. Cascioli et al., *NLO matching for* $t\bar{t}b\bar{b}$ *production with massive b-quarks*, Phys. Lett. B **734** (2014) 210, arXiv: 1309.5912 [hep-ph].

[109] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007, arXiv: 0811.4622 [hep-ph].

[110] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph].

[111] F. Derue, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into* $b\bar{b}$ *in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, ATL-COM-PHYS-2017-079 (Feb. 2017), URL: https://cds.cern.ch/record/2244360.

[112] S. Schumann and F. Krauss,
A Parton shower algorithm based on Catani-Seymour dipole factorisation, JHEP **0803** (2008) 038, arXiv: 0709.1027 [hep-ph].

[113] S. Höche et al., *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph].

[114] C. Anastasiou et al.,
High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO, Phys. Rev. **D69** (2004) 094008, arXiv: hep-ph/0312266 [hep-ph].

[115] S. Alioli et al.,
NLO single-top production matched with shower in POWHEG: s- and t-channel contributions, JHEP **0909** (2009) 111, arXiv: 0907.4076 [hep-ph].

[116] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph].

[117] S. Frixione et al., *Single-Top Production in MC@NLO*, JHEP **03** (2006) 092, arXiv: hep-ph/0512250.

[118] R. Frederix, E. Re, and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, J. High Energy Phys. **1209** (2012) 130.

[119] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, eprint: arXiv:1103.2792.

[120] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. D **82** (2010) 054018, eprint: arXiv:1005.4451.

[121] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, eprint: arXiv:1001.5034.

[122] M. Aaboud et al., *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76**.12 (2016) 653, arXiv: 1608.03953 [hep-ex].

[123] G. Aad et al., *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. **C76**.5 (2016) 292, arXiv: 1603.05598 [hep-ex].

[124] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024 (2016), URL: https://cds.cern.ch/record/2157687.

[125] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-015 (2015), https://cds.cern.ch/record/2037613.

[126] G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv: 1406.0076 [hep-ex].

[127] S. Frixione et al., *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph].

[128] "Multi-Boson Simulation for 13 TeV ATLAS Analyses," tech. rep. ATL-PHYS-PUB-2016-002, CERN, Jan. 2016, URL: https://cds.cern.ch/record/2119986.

[129] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an].

[130] G. Cowan et al., *Erratum to: Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **73** (2013).

[131] ATLAS Collaboration, *Search for the Higgs boson produced in association with a W boson and decaying to four b-quarks via two spin-zero particles in pp collisions at 13 TeV with the ATLAS detector* (2016), arXiv: 1606.08391 [hep-ex].

[132] ATLAS Collaboration, *Direct search for exotic Higgs boson decays to two spin-zero particles in the $H \to 2a \to 4b$ channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATL-COM-PHYS-2015-1457 (Dec. 2015), URL: https://cds.cern.ch/record/2110199.

Appendix A

Fit results for other Signal Mass Hypotheses

This section includes the fit results for the other a-boson mass hypotheses considered, $m_a = 15$ GeV and $m_a = 25$ GeV. The fit results are consistent with those obtained for the $m_a = 20$ GeV mass hypothesis included discussed in Section 9.6.

A.1 Fits to the Asimov dataset

This section shows the signal plus background fits to the Asimov dataset for the regions described previously. In this case, the expected background is fitted to the Asimov dataset (the nuisance parameters are set at 0 while the normalization factors are set at 1, by definition) under the signal-plus-background hypothesis. The fitted regions are shown in Figures A.1 and A.2. Figures A.3 and A.4 shows the fitted nuisance parameters, normalization factors and gamma parameters to the Asimov dataset, as well as the effect of the different nuisance parameters on the parameter of interest.

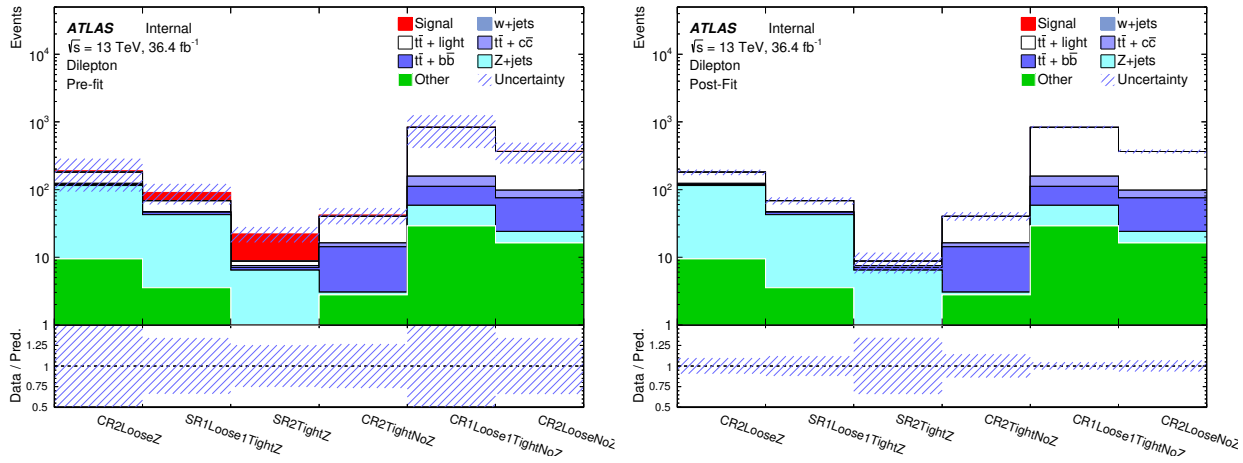

Figure A.1: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 15$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

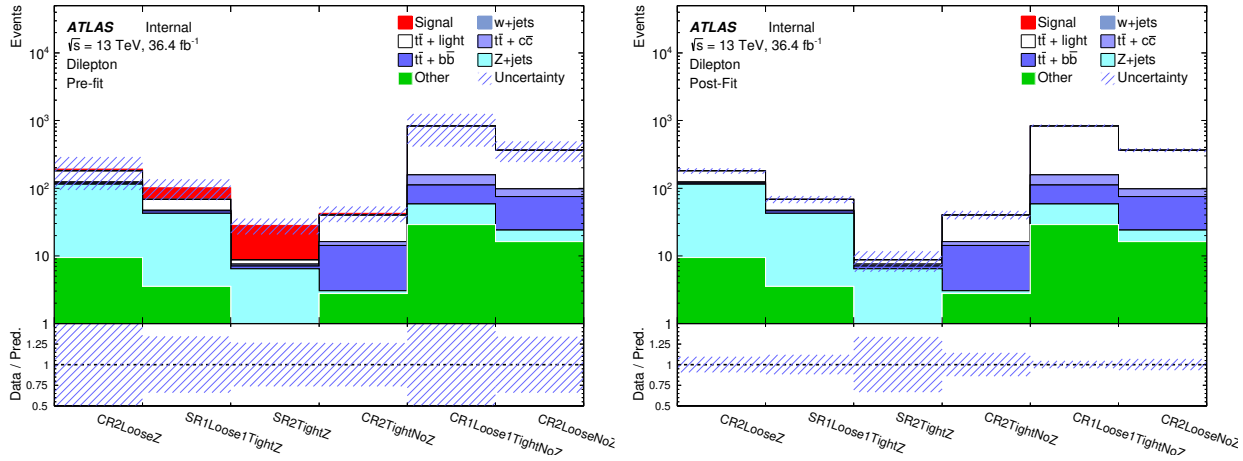

Figure A.2: Signal and control regions showing the predicted yields in simulation (left) before and (right) after the fit to the Asimov dataset. The expected signal yield for the $m_a = 25$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

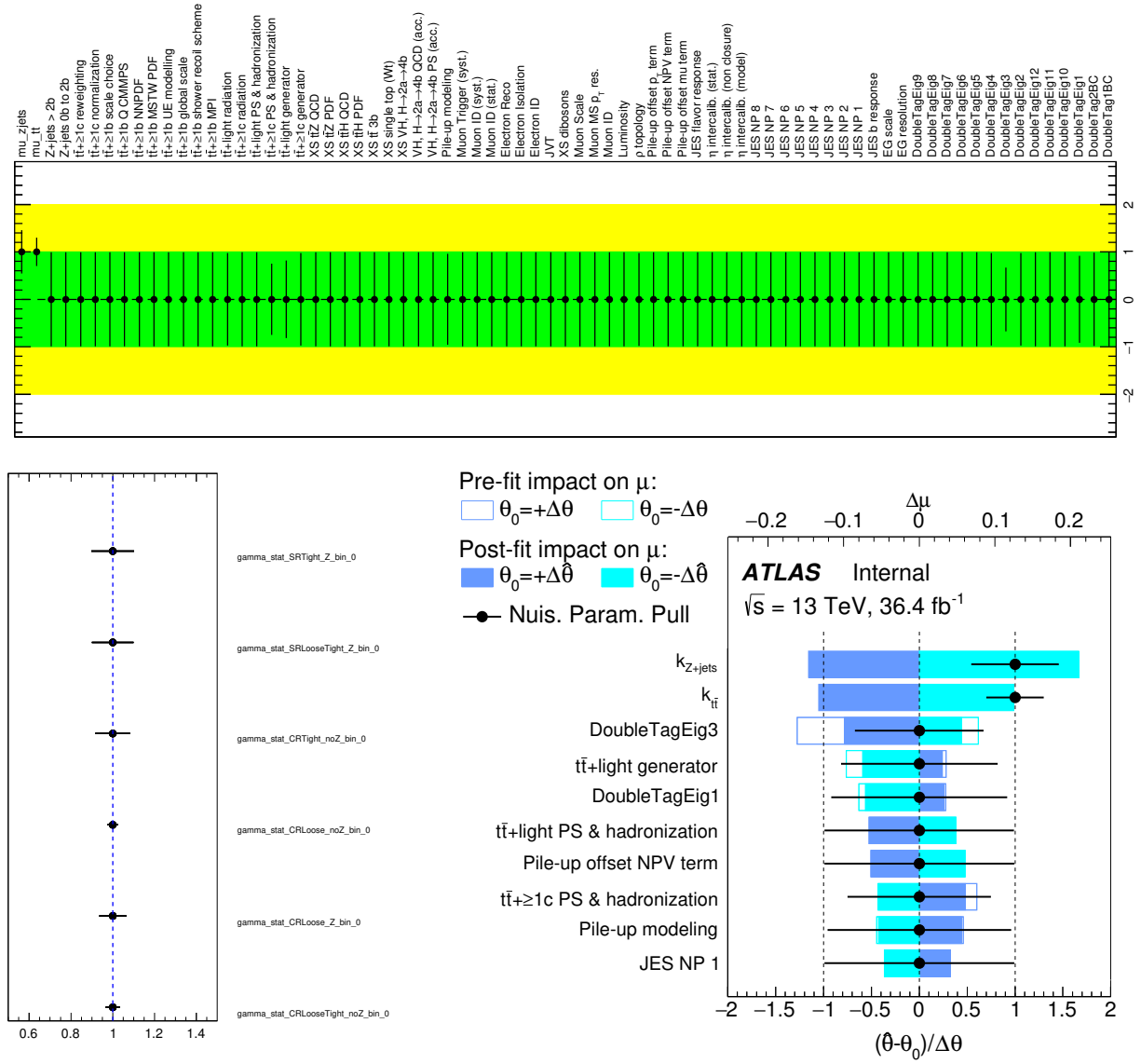

Figure A.3: Nuisance parameters and normalization factors (top), gamma parameters (bottom left) and rankings based on the impact on the signal strength (bottom right) for the Asimov data fit for the signal hypothesis with $m_a = 15$ GeV using the full set of systematics.

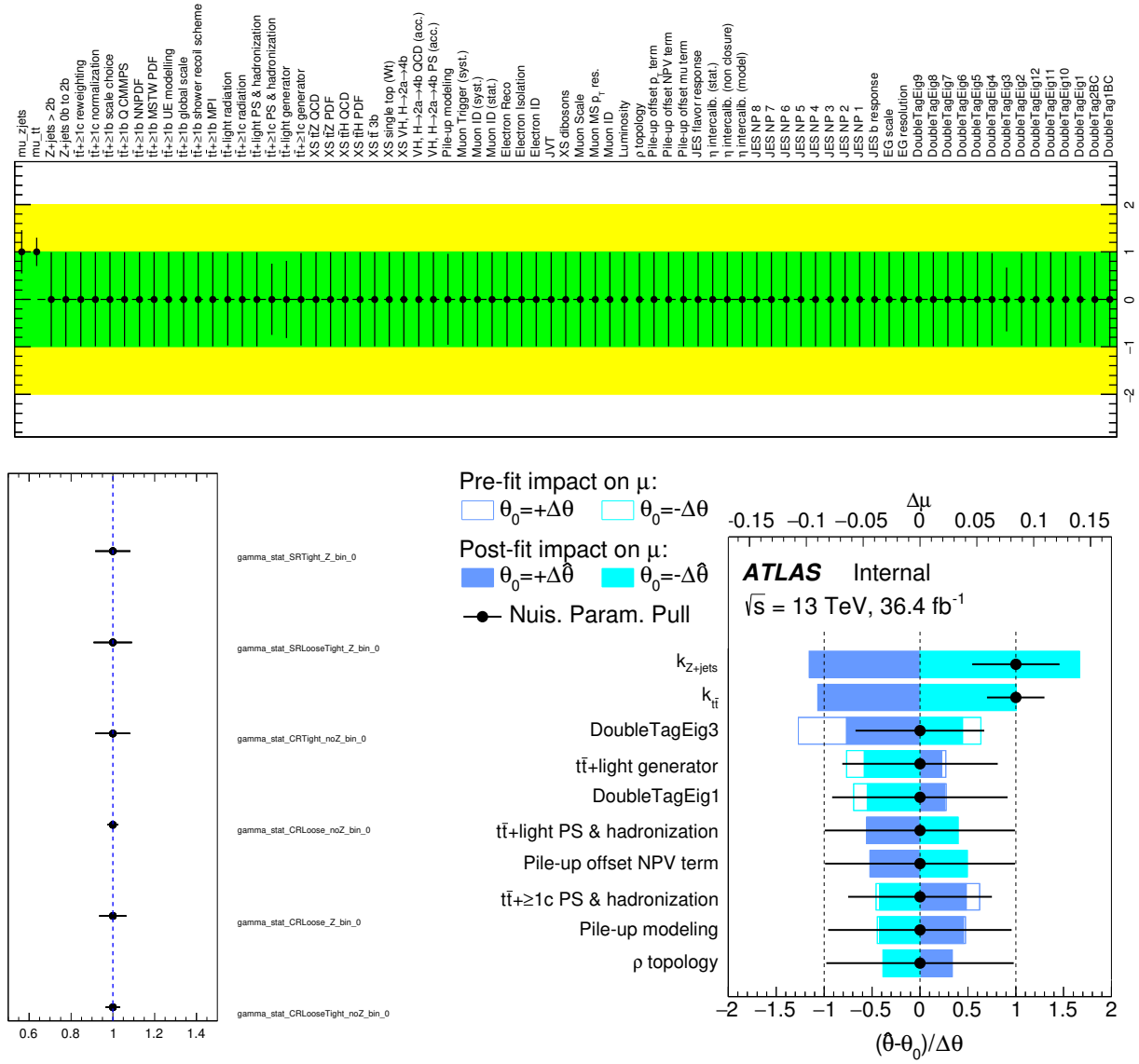

Figure A.4: Nuisance parameters and normalization factors (top), gamma parameters (bottom left) and rankings based on the impact on the signal strength (bottom right) for the Asimov data fit for the signal hypothesis with $m_a = 25$ GeV using the full set of systematics.

A.2 Fits to blinded data

This section shows results for the analysis using full background-only fits to data with the blinded signal regions. Figures A.5 and A.6 show a comparison of the yields in data and simulation for the control regions before and after the blinded fit to data. The yields from simulation in the signal regions are also shown.

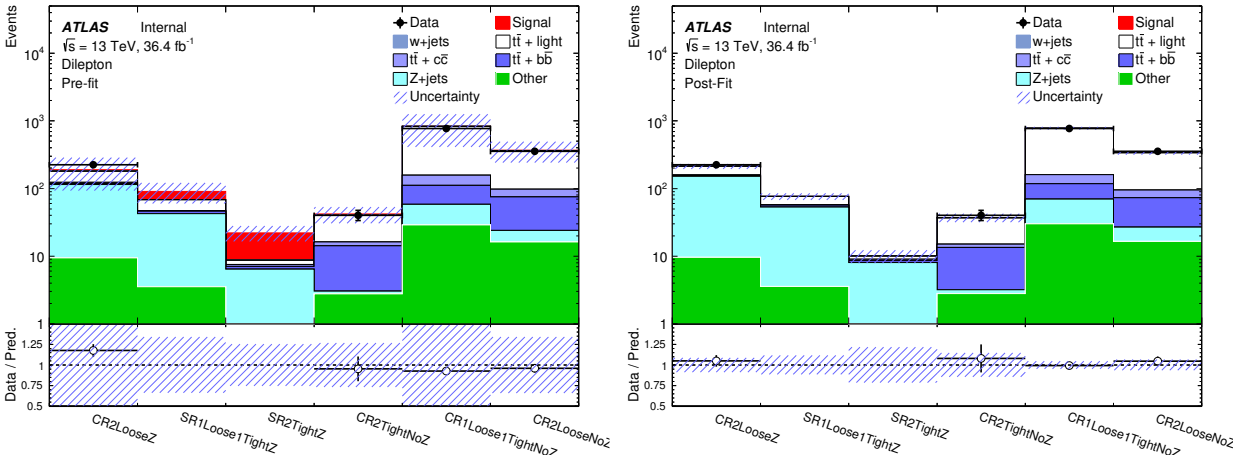

Figure A.5: Predicted yields for the signal and control regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 15$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

The nuisance parameters, as well as the normalization factors and the gamma parameters for the background-only fit are shown in Figures A.7 and A.8. The constraints in the different nuisance parameters are consistent with those found in the Asimov fit, but since these fits are with real data, some pulls are observed.

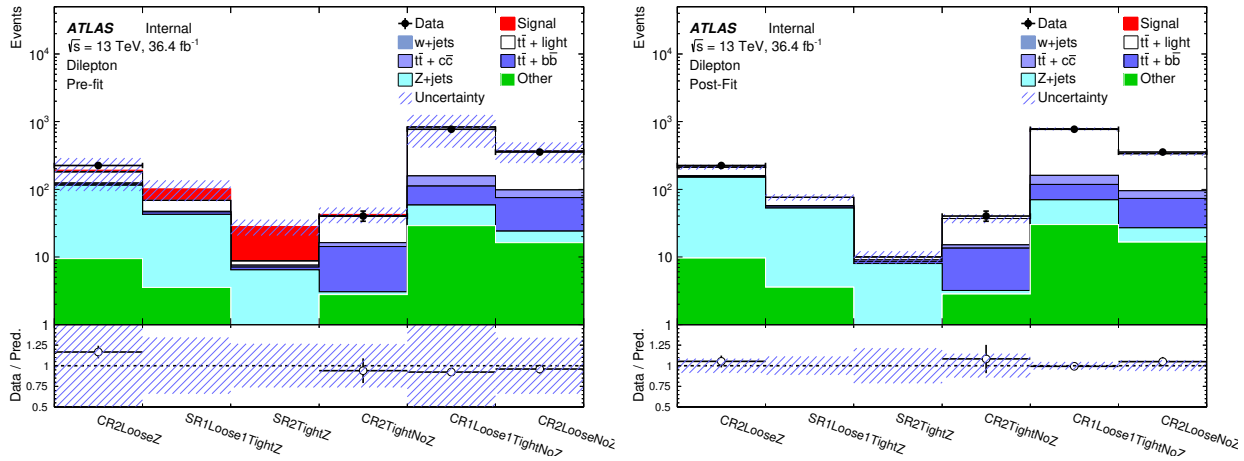

Figure A.6: Predicted yields for the signal and control regions in the blinded fit to data (left) before and (right) after the fit. The data and simulated yields are compared in the control regions, while only the simulated yields are shown in the signal regions. The expected signal yield for the $m_a = 25$ GeV mass hypothesis assuming the SM production cross section for σ_{ZH} and BR$(a \to bb) = 1$ is shown before the fit, while the signal yields set to 0 in the fit (background-only hypothesis).

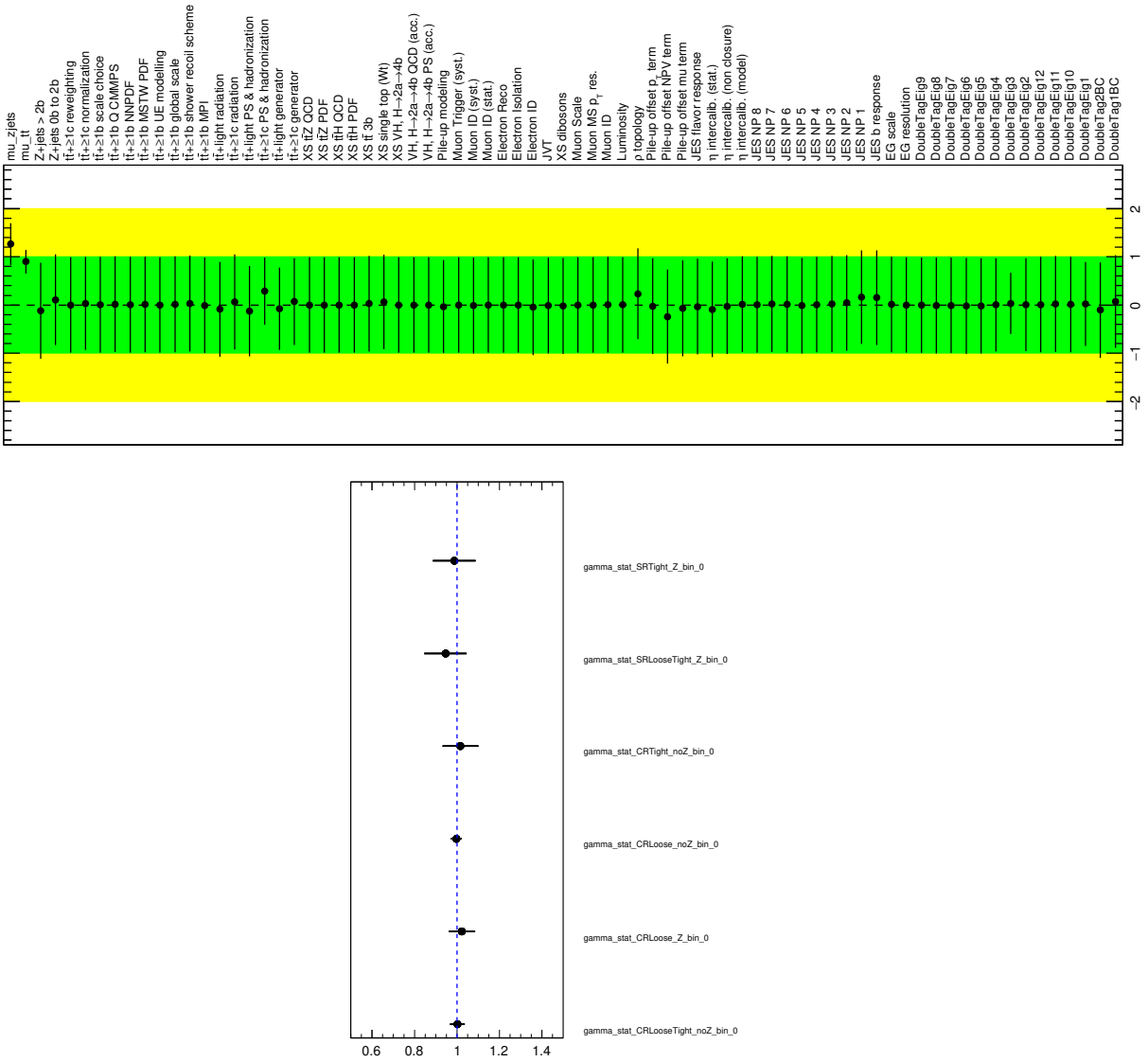

Figure A.7: Nuisance parameters and normalization factors (top), gamma parameters (bottom) for the blinded data fit for the $m_a = 15$ GeV signal hypothesis using the full set of systematics.

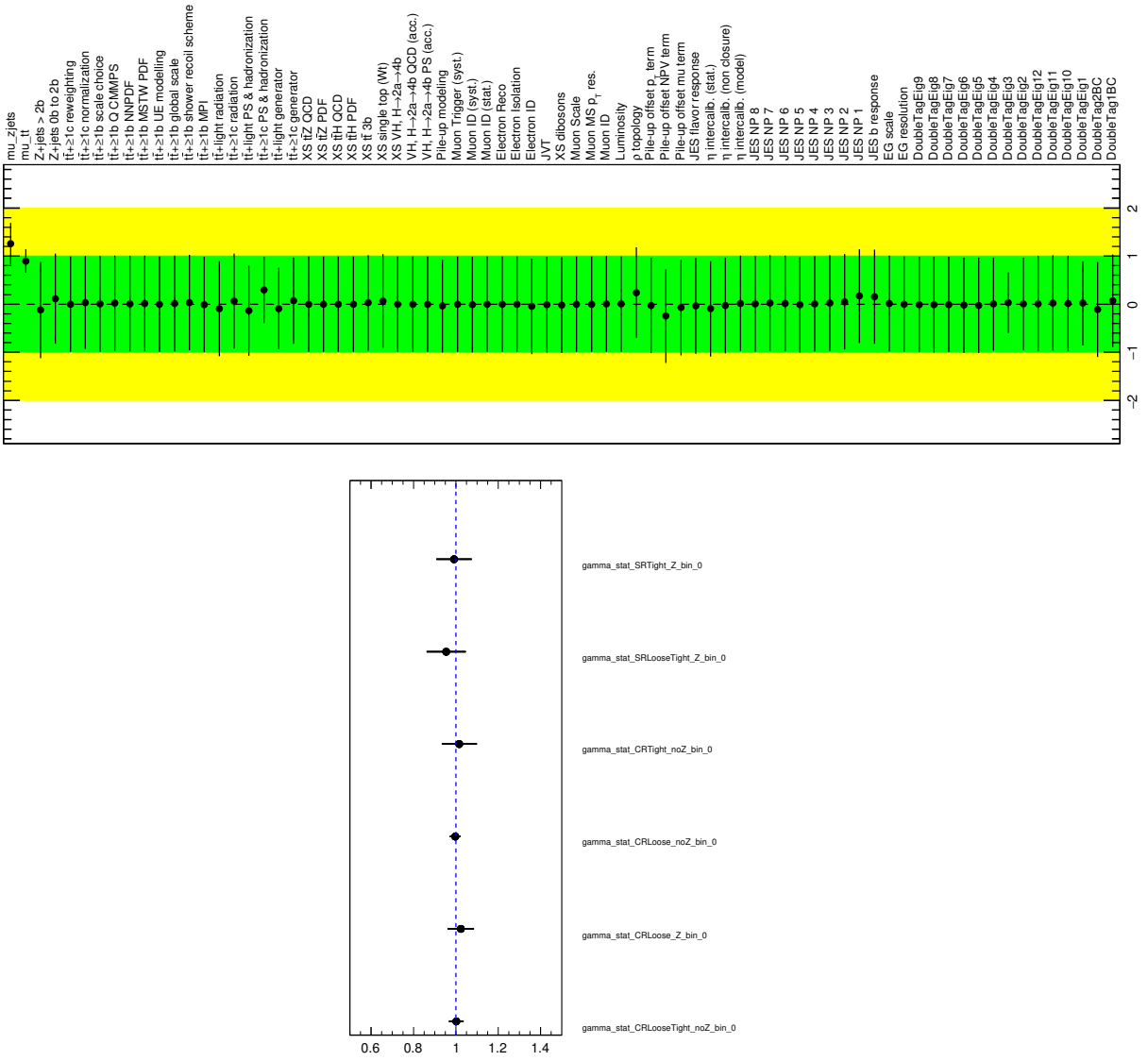

Figure A.8: Nuisance parameters and normalization factors (top), gamma parameters (bottom) for the blinded data fit for the $m_a = 25$ GeV signal hypothesis using the full set of systematics.